



GREEN PATH: an expert system for space planning and design by the generation of human trajectories

Marina Paolanti¹ · Davide Manco² · Rocco Pietrini² · Emanuele Frontoni¹

Received: 14 April 2023 / Revised: 21 September 2023 / Accepted: 8 January 2024 /
Published online: 15 February 2024
© The Author(s) 2024

Abstract

Public space is usually conceived as where people live, perceive, and interact with other people. The environment affects people in several different ways as well. The impact of environmental problems on humans is significant, affecting all human activities, including health and socio-economic development. Thus, there is a need to rethink how space is used. Dealing with the important needs raised by climate emergency, pandemic and digitization, the contributions of this paper consist in the creation of opportunities for developing generative approaches to space design and utilization. It is proposed GREEN PATH, an intelligent expert system for space planning. GREEN PATH uses human trajectories and deep learning methods to analyse and understand human behaviour for offering insights to layout designers. In particular, a Generative Adversarial Imitation Learning (GAIL) framework hybridised with classical reinforcement learning methods is proposed. An example of the classical reinforcement learning method used is continuous penalties, which allow us to model the shape of the trajectories and insert a bias, which is necessary for the generation, into the training. The structure of the framework and the formalisation of the problem to be solved allow for the evaluation of the results in terms of generation and prediction. The use case is a chosen retail domain that will serve as a demonstrator for optimising the layout environment and improving the shopping experience. Experiments were assessed on shoppers' trajectories obtained from four different stores, considering two years.

Keywords Generative adversarial imitation learning · Trajectory generation · Reinforcement learning · GANs

✉ Rocco Pietrini
r.pietrini@staff.univpm.it

Marina Paolanti
marina.paolanti@unimc.it

¹ University of Macerata, Department of Political Sciences Communication and International Relations, 62100 Macerata, Italy

² Università Politecnica delle Marche, Department of Information Engineering (DII), 60131 Ancona, Italy

Abbreviations

Acronym	Description
ADE	Average Displacement Error
ADTW	Average Dynamic Time Warping
AI	Artificial Intelligence
ALCSS	Average Longest Common Subsequence
BC	Behavioural Cloning
CDF-AUC	Cumulative Distribution Function Area Under the Curve
CST	Cumulative Stopping Time
DCB	Dynamic Context Beliefs
DTW	Dynamic Time Warping
FoWM	Fog of War Map
GAE	General Advantage Estimation
GAN	Generative Adversarial Network
GAIL	Generative Adversarial Imitation Learning
HBA	Human Behaviour Analysis
HMM	Hidden Markov Model
ICT	Information and Communication Technology
IRL	Inverse Reinforcement Learning
KPI	Key Performance Indicator
LCSS	Longest Common Subsequence
LSTM	Long Short-Term Memory
RMSProp	Root Mean Square Propagation
RC-GAN	Recurrent Conditional Generative Adversarial Network
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
RTLS	Real-Time Location System
STC	Similar Trajectories Count
TP	Trajectory Prediction
UWB	Ultra-Wideband
WGAN	Wasserstein Generative Adversarial Network

1 Introduction

The convergence of heightened focus on quality of life, well-being, and societal climate resilience, alongside the transformative impact of the ongoing global COVID-19 pandemic, has created an opportunity for us to reconsider how we plan and design spaces. Besides, architectural design features such as shape or building orientation have a significant influence on energy loads and their trade-offs [26]. This creates various new challenges to re-imagine the way of designing, operating, and using these common areas, thus bringing attention to analysing the behaviour of people who interact and live in public spaces. Tuan [41] states that “place is space infused with human meaning” and argues for two important concepts that humans are rooted in place and possess and cultivate a sense of place. The impact of environmental problems on humans is significant, affecting all human activities, including health and socio-economic development. Thus, there is a need to rethink how space is used. A well-designed public space matches the multiple needs of everyday and one-time users and it should respect the European guidelines that advocate sustainable issues for planning.

Providing planners with Information and Communication Technologies (ICT) tools that can facilitate the definition of guidelines or protocols for their investigation should be fundamental to determining the significance of such areas for society [32].

1.1 Challenges

Given such a premise, and considering the recent literature in the field, the design of public space starts from what occurs in that environment. Understanding the behaviours, actions and attitudes of people living in that space outperforms the standard rules of design that usually are applied. Human Behaviour Analysis (HBA) has positive outcomes from the prediction, generation, and simulation of human behaviour [19]. Human trajectory is a research area in HBA; when humans move in a given environment, they intuitively follow unwritten social rules [34]. Their behaviour also strongly depends on the type of environment in which they operate (e.g., malls, parks, or sidewalks). Predicting human behaviour by trajectory prediction is a burdensome task in several aspects. Recent research on computer vision has addressed these challenges by overcoming these limitations. Among them, the most important are the following [12]:

- *Socially acceptable movements.* Some paths are physically possible but are usually not performed to comply with implicit social rules, such as respecting minimum interpersonal distance.
- *Human-space interaction.* The surrounding environment affects human actions. Obstacles and objects all have, in one way or another, an effect on human behaviour. It is, therefore, important to model these interactions and try to describe them.
- *Human-human interaction.* Human trajectories depend heavily on how the people around them behave. A human being can predict the behaviour of other people and, consequently, make movements to avoid them.
- *Multimodality.* In HBA, since human behaviour is unpredictable, the possible behaviours are several and the correct solutions are different.
- *Generalisability.* A method should be evaluated for its ability to predict the entire distribution of possible human trajectories.

The solution to this problem concerns many practical applications, ranging from data visualisation to simulation applications. It is possible to deduce if a specific configuration of the environment models human behaviour and how it influences Key Performance Indicators (KPIs). If there is an environment with physical, social or semantic limits and constraints, it is possible to correctly predict or simulate the flow of human trajectories for a specific period and generalisable mode. State-of-the-art approaches are primarily based on the use of Generative Adversarial Neural Networks (GANs) or Long Short Term Memory (LSTM) [4], [12] in crowded environments [30], and most of them do not model user behaviour in the surrounding environment but merely generate acceptable and realistic trajectories. In [34], we have filled this gap by proposing and defining new methods and metrics to help understand trajectories. In particular, new deep learning models based on LSTM and GAN architectures are used in both unimodal and multimodal contexts.

However, frameworks based on Inverse Reinforcement Learning (IRL) closely approximate trajectories produced by humans [17], [46], and Generative Adversarial Imitation Learning (GAIL) is proven to be a powerful and practical approach for learning sequential decision-making policies [13]. GAIL allows us to find a correlation between objects present

in the scene and proximity to the search target. It is possible to find an analogy between the search for a certain category of products inside the retail space and the correlation of these products with others close by.

1.2 Nature and scope

In this regard, this paper aims to present GREEN PATH (GREEN space Planning by prediction and generAtion of Trajectories of Humans) a system for the creation of opportunities for developing resilient and regenerative approaches to public space design and utilization. The goal of GREEN PATH is not merely the design of a space, but also the creation of a new model more sustainable, more agile, and smarter and can generate human trajectories in an environment with complex constraints. GREEN PATH uses human trajectories and deep learning methods to analyse and understand human behaviour for offering insights to layout designers.

In this regard, the paper aims to propose a predictive and generative model that can handle an environment with complex constraints. In particular, it proposes a framework based on the work of [46] hybridised with classical reinforcement learning methods, such as continuous penalties, which allow for modelling the shape of the trajectories and inserting a bias in the training necessary for the generation. The structure of the framework and the formalisation of the problem to be solved allow for the evaluation of the results in two aspects: prediction and generation. Generation refers to the creation of trajectories from scratch, with determined points of origin. These trajectories are completely new, and the evaluation is done on quality and efficiency. The efficiency is similar to human efficiency, while the quality indicates the ability to create realistic trajectories and is evaluated by comparing the generated trajectories with those of the test set. Forecasting refers to the prediction of future paths for real trajectories that have already started. From this point of view, the geometric proximity of the generated points to the real ones is verified.

The approach has been applied to real scenarios, and the experiments were assessed on four datasets derived from different stores over two years. The behaviour of 10.4 million visitors was analysed, as described in [8, 29].

1.3 Contributions

GREEN PATH will make extensive use of AI for automatizing i) human behaviour understanding and forecasting through the creation of a widely generalizable system that allows to generation of human trajectories trajectories from zero, ii) space interpretation and virtualization, in fact, a representation of the state that can be easily expanded to different contexts. Largely inspired by the Dynamic Context Beliefs (DCB) of Yang et al. [46] and taking inspiration from the videogame world, a dynamic representation system has been developed iii) content creation and human-space interaction with the verification and resolution of the problem concerning the form of the reward existing in the work carried out by Yang et al. [46], in particular as regards the formulation of the reward function. Finally, iv) design and arrangement. While a manually tagged dataset is used for training, this does not exclude that the source of the states may be different. Since the state is based on exploration, it is possible to generalize its creation during the deployment phase and also carry it out through other methods, such as visual input from a robot.

The paper is organised as follows. Section 2 provides an overview of state-of-the-art approaches for trajectory prediction and generation. Section 3 presents the proposed

approach, which is based on GAIL. Section 4 presents a comparison between our approach and several state-of-the-art algorithms, along with a detailed analysis of our framework. Limitations are presented in Section 5. In Section 6, conclusions and discussions are presented while future directions for this research have been proposed in the last Section 7.

2 Related works

Human trajectories are information-rich features that can help in understanding the environment, giving an idea about the interactions between objects and ongoing events [28]. Modelling human behaviour has overwhelming potential, especially from an economic and strategic point of view. When people walk in a space, they adhere to a huge number of unwritten trivial rules and observe social practices [20]. For instance, if they move inside a space, they respect their paths and yield to other nearby people to have their right of way. The competence to model these unwritten rules and apply them to predict, understand and generate users' movements in an environment is extremely worthwhile for the design of intelligent tracking systems in smart environments. This problem appears challenging since several issues arise for the prediction and generation of human actions, taking into account such common sense behaviour [1].

Tracking people to understand human behaviour has a long tradition in computer vision literature [31], [21], [45], [39]. However, recently, predictive models have gained increased interest [44], [43]. Trajectory prediction is achieved by modelling and learning human-space [3], [14] or human-human interactions [33], [24].

Predictive models of pedestrian dynamics have been developed by encoding the coupled nature of multi-pedestrian interactions using game theory and deep-learning-based visual analysis to estimate person-specific behaviour parameters [24], [22]. In particular, the authors used concepts from game theory to model the intertwined decision-making processes of multiple pedestrians. Moreover, they used visual classifiers to learn a mapping from pedestrian appearance to behaviour parameters.

Social acceptability has been inspected using data-driven techniques based on Recurrent Neural Networks (RNNs). [1] proposed a model called Social LSTM, which can learn general human movement and predict future trajectories. The proposed model can simultaneously predict the paths of all the people in a scene, considering the common sense rules and social conventions that humans generally adopt as they operate in public environments. In particular, the author introduced a “social” pooling layer that allows LSTMs of spatially proximal sequences to share their hidden states.

Bartoli et al. [4] extended the work of Alahi et al. [1] by defining “context-aware” pooling that allows the model to deal with static objects in the region around a person. In particular, their approach is based on the LSTM network that can learn and predict human movement in crowded environments.

To address the limitations of the aforementioned works, Gupta et al. [12] exploited GANs to generate multiple socially acceptable trajectories, given an observed past. These behaviours concern socially accepted motion trajectories in crowded spaces. Their model is called “Social GAN” since they addressed the multimodality of trajectories.

Kothari et al. [20] define trajectory predictions as “given the past trajectories of all humans in a scene, forecast the future trajectories which conform to the social norms”. To focus on learning the social interactions that affect human motion, the authors assume that there do not exist any physical constraints in the scenes. They also focus on short-term human trajectory

forecasting (next 5 secs). [47] presents early experimental results obtained including social information in their convolutional model using occupancy grids and maps. These experiments empirically showed that occupancy methods are ineffective in representing social information and did not improve their results.

These works are milestones for human-human interactions. Moreover, their purpose is to predict micro-trajectories, i.e. the precise generation of points following the current one. While the interest of this work is mainly related to macro-trajectories, as stated in the Introduction section, this paper also focuses on multimodality and human-space interaction.

In this regard, Kim et al. [17] proposed a framework for socially adaptive path planning in dynamic environments. In particular, they used an IRL module that adopted a set of trajectories generated by an expert for learning expert behaviour with several state features.

In [46], the authors proposed the first IRL model for learning the internal reward function and policy used by humans during a visual search. The purpose of this work was to reproduce the trajectory of the human gaze as it searches for a given object within the image. The theoretical basis of this work is the association that the human mind makes between objects that are necessary for or related to the achievement of a given goal.

A bottleneck of reinforcement learning is that it concerns the optimisation of a predefined reward function [38]. The design of a suitable reward function can be arduous in complex environments. Imitation learning approaches have proven to close this gap by learning how to perform tasks directly from expert demonstrations [15]. Among these, GAIL is a model-free imitation learning method that is highly efficient [13].

In this context, Li et al. [23] proposed an algorithm that can deduce the latent structure of expert demonstrations in an unsupervised manner. Their method, based on GAIL, can not only emulate complex behaviours but also learn interpretable and essential representations of behavioural data as visual demonstrations. The domain of application was autonomous driving to mimic human behaviours related to driving a vehicle. The results obtained were fair, despite the difficulty of the task. The most interesting part of this work was the improvements to GAIL performance using a modified version of the Wasserstein loss [2], often used on GANs since it allows to eliminate some problems, such as vanishing gradients or the possibility of getting stuck in a minimum location. They also used “reward augmentation” [6], which consists of adding an a priori reward, which models a bias to be reflected in the model training, to the reward provided by the discriminator [6].

The work of Ferracuti et al. [7] concerns the retail environment and uses Real-Time Locating System (RTLS) tags to collect human trajectory data. The tags were used to infer visitors’ preferred paths and their segmentation. In the same context, Paolanti et al. [28] presented a smart mechatronic system (sCREEN, Consumer REtail ExperieNce) for indoor navigation assistance. The system is based on a new Hidden Markov Model (HMM) to represent shoppers’ shelf/category attraction and usual retail scenarios (shelf-out-of-stock and modification of store layout).

In [5], the authors proposed a unified deep learning framework for the generation and analysis of driving scenario trajectories and validated its effectiveness in a principled way. To model and generate scenarios of trajectories with different lengths, they have developed two approaches. Firstly, they adapted a Recurrent Conditional Generative Adversarial Network (RC-GAN) by conditioning the length of the trajectories. Then, they designed an architecture based on a Recurrent Autoencoder with GANs to obviate the variable length issue, wherein they trained a GAN to learn/generate the latent representations of original trajectories.

Based on the idea proposed by [46], this paper attempts to solve the generation of trajectories inside a store. Assuming that the elements of the scene in the case of gaze prediction are similar to the categories near the customer in the case of movements inside a store, it is

also possible to generalise the forecasting of trajectories starting from paths already partially formed. In this way, it is possible to understand, for example, the path taken by a customer starting from any point to go to each of the categories in the store. It also allows the management of different cases and trajectories in such a way as to foresee any possible behaviour and movement of the customer, by relying on statistical data for an estimate of the probability of purchase or interest in each category. Marketing strategies can be developed for specific customers, which can also be applied in real-time.

3 Materials and methods

GREEN PATH exploits an advanced intelligence system made of RTLS tags to collect human trajectory data, which, analyses, monitors, and understands everything that happens inside the target area. GREEN PATH provides alternatives in the design process. Following the idea presented in [46], we propose a GAIL framework to predict human trajectories in real environments. The framework aims to model such behaviour by a state representation that considers the influence of the environment on the short-term decision-making process of the user. A sparse matrix is adopted, which comprises C channels and has a fixed dimension of 47×47 . Every channel contains a representation of the position of a certain category in the target store. Our dataset consists of data collected from four stores that are encoded in this way. For each location, we have several points related to the tags. These tags are placed on a shopping cart or a basket; hence, their points need to be split into trajectories. The framework is comprehensively evaluated on the “Shopper trajectories dataset”, a publicly available dataset. The overall framework of GREEN PATH is depicted in Fig. 1.

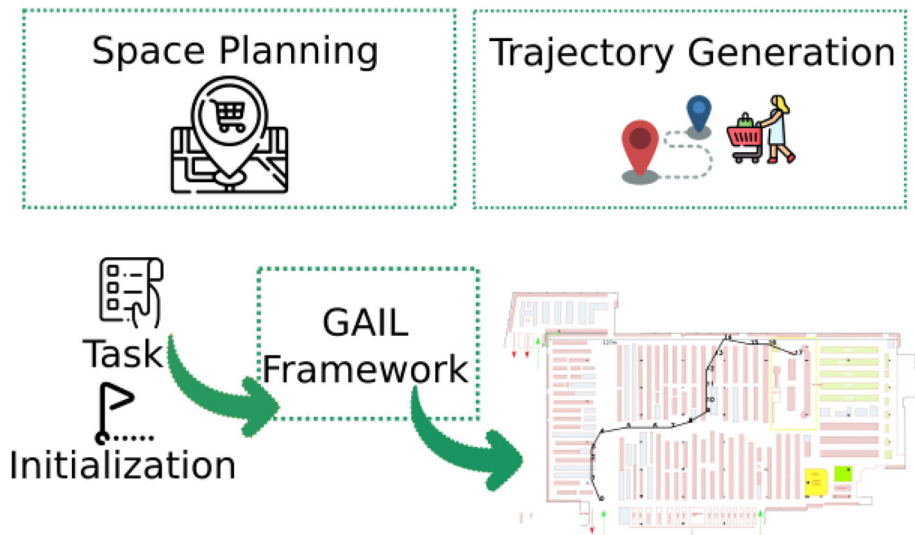


Fig. 1 GREEN PATH workflow

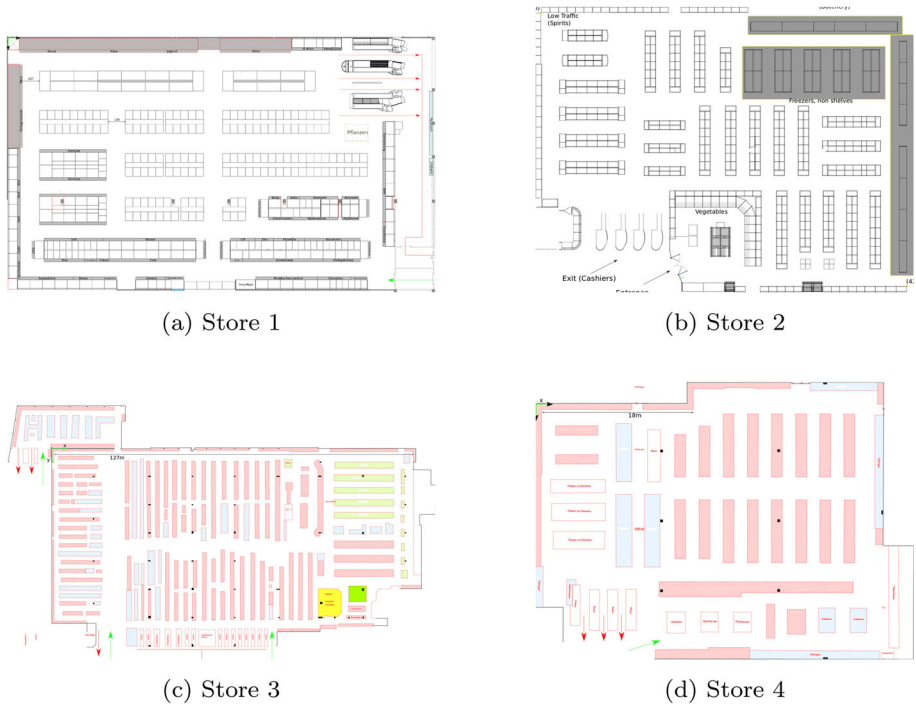


Fig. 2 Layout of the four stores for which the “Shopper trajectories dataset” has been collected

3.1 Shopper trajectories dataset

The dataset used in this work was acquired from four different stores in Germany and Indonesia, measuring the behaviour of 10.4 million shoppers over two years, as described in [8]. The data were collected with a tracking system based on Ultra Wideband (UWB) technology, with tags embedded in shopping carts. The UWB is suitable for applications where positioning accuracy is a critical issue [7]. This technology uses some UWB antennas that are suitably placed in a fixed area and battery-powered tags that can freely move in the area [28]. Figure 2 represents the layout of the four stores for which the “Shopper trajectories dataset” has been collected.

Table 1 reports the number of data points for each dataset. The number of trajectories in a dataset is approximately proportional to the number of data points.

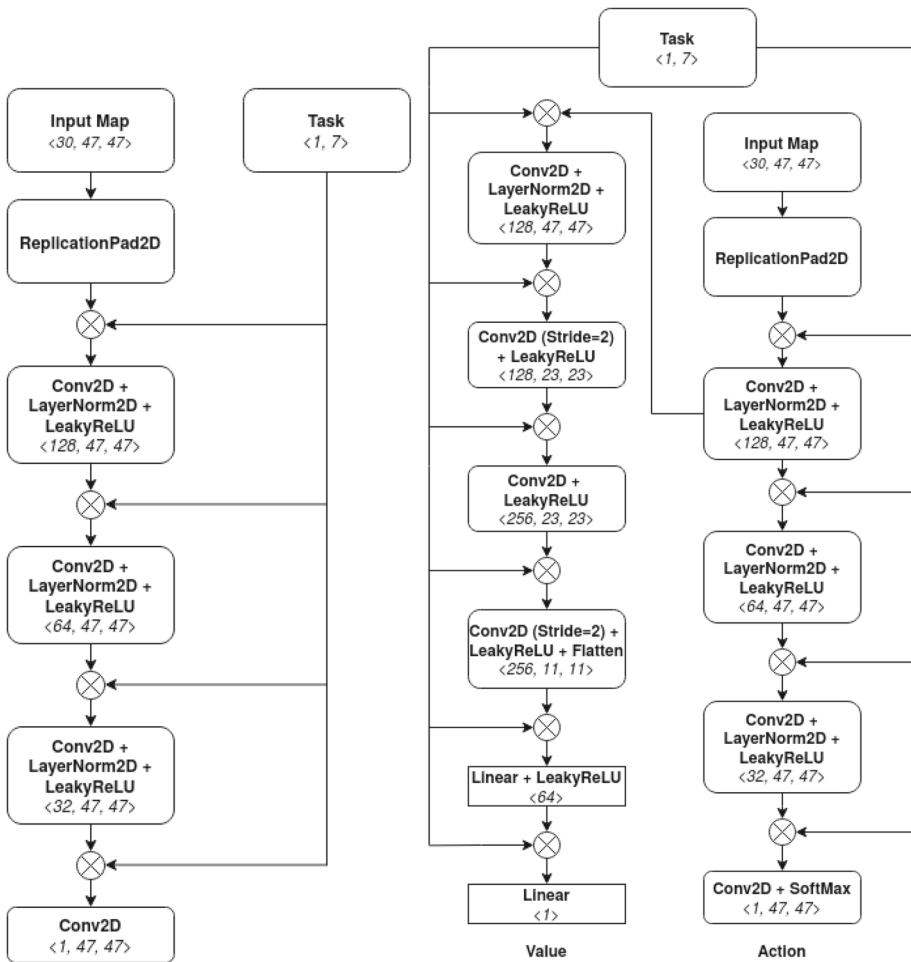
Table 1 Datasets and number of data points

	Total data points
Store1	1,867,920
Store2	3,945,842
Store3	6,416,000
Store4	212,341

3.2 GAIL framework

The GAIL framework [13] is an imitation learning approach similar to inverse reinforcement learning but formally different since it does not explicitly attempt to recover the reward function. In this case, the reward function created is different from the implicit and hidden functions of the expert. The intuition is to create a “judge” (discriminator) that indicates to the agent what he should and should not do based on the data obtained from an expert. The reward increases the more the agent approaches what the judge deems correct.

Our framework (see Fig. 3) consisted of three networks: the discriminator, the agent (the generator) and the critic. The critic and agent networks shared one layer of feature extraction. The discriminator and the agent have an identical structure, unlike the Scanpath Prediction case, where there were some small differences in padding and kernel size. There were four



(a) Discriminator.

(b) Generator and critic.

Fig. 3 Our GAIL framework

layers of 128, 64, 32 and 1 filters, respectively, all with a kernel of 3×3 size and zero-padding of 1, except the first layer, which will have a reflection padding. This choice was experimentally dictated by the presence of higher probabilities (for the generator) and rewards (for the discriminator) along all the edges, which implies that zero-padding for the padding of the states is not the most suitable choice. The critic will share the first layer with the generator network. In this work, some convolution layers were added, with the same dimensions as the previous layer to be downsampled but with a stride of 2, to reduce the size of the output maps. This choice allowed better effectiveness of downsampling since it allows the network to learn additional parameters to increase generalisation without eliminating information with a predefined method, such as the maximum function. This approach is often recommended when using GANs, even if it has a greater number of parameters to train. The three networks used a dropout of 0.2 on most layers, and the activation function from Rectified Linear Unit (ReLU) to Leaky ReLU was replaced to avoid the problem of dying ReLU. The generator and critic networks were initialized through a Xavier (also called “Glorot”) initialisation [9], while the standard PyTorch initialisation was preferred for the discriminator since it would still be effective. Root Mean Squared Propagation (RMSProp) was chosen as the optimiser for training the discriminator, a choice derived for reasons of greater effectiveness demonstrated in cases of loss that used gradient penalties [11, 25]. Finally, regarding the loss, tests were carried out with the Wasserstein Generative Adversarial Network (WGAN) version with gradient penalty and with the normal GAN version, which also has a gradient penalty centred at 0, and applied only to real data, since the literature guarantees better convergence and generalisation using this method [16, 25, 35]. We concatenated the chosen task and the output on every layer, as already done in [46]. In this way, we obtained a correlation between the chosen task and the action taken by the agent. In Fig. 3, it can also be seen that $C = 30$, whereas, there are only seven tasks. Not all categories were considered for the possible tasks, as many of them did not have that many customers with considerable Cumulative Stopping Time (CST). C also includes the Fog of War Map (FoWM), detailed later, and a further map that contains the previous positions of the agent. The actor and critic networks were trained using proximal policy optimisation [37], with a learning rate of 0.00001 and a discount factor of 0.9, and advantages were estimated using General Advantage Estimation (GAE) [36]. Other hyperparameters were set to be equal to the original paper’s suggested values. The discriminator was trained using the standard GAN loss [10], with a learning rate of 0.00005.

3.2.1 Preprocessing

To obtain good results, it is mandatory to choose a proper splitting strategy that correctly models our requirements. The goal was a generalised framework that, given a certain map and a target category, generates a user trajectory towards that category. Therefore, the chosen splitting strategy should be related to an inferred task of the customers. To infer such tasks, we evaluated a CST, i.e. the time during which the points of a tag were stationary near a certain category. Then, we split the trajectory when the CST exceeded a certain threshold. We chose the last stopping point as the last point of the trajectory. We then initialised another trajectory by using the same ending point of the previous trajectory as a starting position. In this way, we also obtained a good generalisation for the generated trajectories, as they did not depend on the initialisation. We also split the trajectories when they reached a selected entrance or exit area. Lastly, after doing this separation, we had to filter and discretise these points in a 47×47 grid. To decrease the number of points, we also used sampling by considering only

points that have at least a Euclidean distance of 3 (calculated on the grid), but not greater than 5, with the previous point on the trajectory.

We formalised the environment as a Markov decision process; therefore, we had to define our trajectories as a set of state-action couples. In our case, the state was the portion of the store that the customer had seen so far, which was expressed as a map with zeroes in all the positions. We were inspired by video-games and the military concept of the “fog of war”. Hence, we dynamically upgraded the current state by following the exploration of the agent. If the agent decides to move to a certain location, the current state moves to a new state that has actual map values within a radial area from the new position. At the k -th step, we have a cumulative map that has the real categories’ values on the explored area and zeroes elsewhere (the “fog of war”). To further augment the information of the agent, we added a FoWM as a new channel in the current states matrix. This map has ones at the points that were not explored by the customer.

3.2.2 Reward augmentation

The action space is also a 47×47 matrix; hence, the agent can theoretically go everywhere on the map at every step. We have experimentally observed that using a hard constraint on the movement of the agent leads to a bad convergence of the overall framework. Therefore, we chose to adopt a “reward augmentation” method, as in [23], that models a soft constraint. We penalised the agent if he chose to move to a point farther than a preset ϕ radius. Thus, we assigned him a penalty formulated as (1):

$$P_0 = \max(0, \lambda_{P_0}(\sqrt{(x_{i-1} - x_i)^2 + (y_{i-1} - y_i)^2} - \phi)) \quad (1)$$

This penalty is always non-negative and is subtracted from the final reward. The parameter λ_{P_0} is a constant that controls the influence of the penalty; we set it to a value of 0.1, and ϕ was set to 5. We can also apply similar penalties for movements that are too close to the current position (see (2)):

$$P_1 = \max(0, \lambda_{P_1}(\phi_{Near} - \sqrt{(x_{i-1} - x_i)^2 + (y_{i-1} - y_i)^2})) \quad (2)$$

In our configuration, λ_{P_1} was 0.1 and ϕ_{Near} was 1. It should be noted that these two penalties did not add a real bias in the training, as our preprocessing already sampled for points that were at a distance between 3 and 5 from the previous points. Hence, they were more similar to the “reward shaping” proposed by Ng et al. [27] than the “reward augmentation” of Li et al. [23]. However, we noticed that the dataset had a lot of noise and points that were often located on shelves or walls. Thus, we added a biasing penalty that discouraged the agent from moving towards these “un-walkable” points. This penalty cannot be directly formalised as a non-constant and convex function, but we can use the distance between the current position and the “un-walkable” point as a value multiplied by a parameter λ_{P_2} that we set to 0.1. Another biasing penalty added to improve the linearity of the generated trajectory (and to make the framework more resilient to noise) was a penalty applied to the maximum angle of movement. Consider three points on the trajectory: $p_1(x_1, y_1)$, $p_2(x_2, y_2)$ and $p_3(x_3, y_3)$, with p_3 being the movement that the agent desires to make. We calculate the angle θ_0 as formulated in (3):

$$\theta_0 = \arctan2(\Delta x_{\theta_0}, \Delta y_{\theta_0}) \quad (3)$$

Δx_{θ_0} denotes the difference between x_1 and x_2 , Δy_{θ_0} is the difference between y_1 and y_2 . We calculate θ_1 as follows in (4):

$$\theta_1 = \arctan2(\Delta x_{\theta_1}, \Delta y_{\theta_1}) \quad (4)$$

Like before, Δx_{θ_1} is the difference between x_2 and x_3 , and Δy_{θ_1} is the difference between y_2 e y_3 . In (5), the absolute value:

$$\Delta\theta = |\theta_1 - \theta_0| \quad (5)$$

We will have the angle between the two lines. This angle is the angle variation of the last movement. We convert it to degrees, and if we obtain a value greater or equal to 180, we will use the complementary angle since we need the inner angle. We express the final penalty as (6):

$$P_2 = \max(0, \lambda_{P_3}(\Delta\theta - \theta_{Max})) \quad (6)$$

Where λ_{P_3} was set to 0.01 because this penalty was more biasing than the others. All these penalties hybridised our method with pure reinforcement learning. The final reward will be the difference between the discriminator's result and the penalties.

4 Results and discussions

In the following section, for the evaluation of the obtained results, two different aspects were examined. The first goal was to achieve a search efficiency similar to human efficiency. Our task was to imitate human behaviour and not reach the task zone in the fewest number of steps. Therefore, our reward formulation must adhere to this requirement. Yang et al. [46] used the logarithm of the sigmoid function applied to the result of the discriminator. This function produces only non-positive values; hence, it is trivial to infer that the agent performing the actions will be encouraged to complete the task as soon as possible to minimise the total cost of the trajectory. For the generation, we will use the Target Fixation Probability AUC in the same way as Yang et al. [46]. For this metric, we will only need to change the name of the curve since in our domain, the term "Fixation" loses its meaning, and we will simply refer to it as the Cumulative Distribution Function AUC or CDF-AUC. We will compare the results with those obtained by splitting the test set into two parts and using the first part as if they were generated trajectories. This allows us to have a ground truth to which we aim to approximate or, in the case of metrics related to quality, even surpass. We will refer to the results obtained using this method as "Human." Surpassing the "Human" results in terms of quality metrics does not mean moving away from good imitation. Quality metrics measure the similarity between the generated trajectories and the test set trajectories, indicating a measure of generalization. The "Human" value is only a reference point that, depending on the extracted trajectories, may not be optimal. As the lower limit, we will use trajectories generated with the untrained framework, resulting in completely random points. The values obtained in this way, which we will call "Random," represent the values we aim to distance ourselves from. Finally, we will use the network trained through Behavioural Cloning (BC) as the last reference to evaluate the difference achieved with this method. The trajectories generated with BC will all have real initialization. We call our method Trajectory Prediction (TP).

Figure 4 represents this behaviour using the cumulative distribution function curve. For each step, the relative cumulative probability of reaching the target can be obtained. We compared our method with the efficiency of the human trajectories in the test set and with an untrained generator that created random trajectories. To have another imitation learning

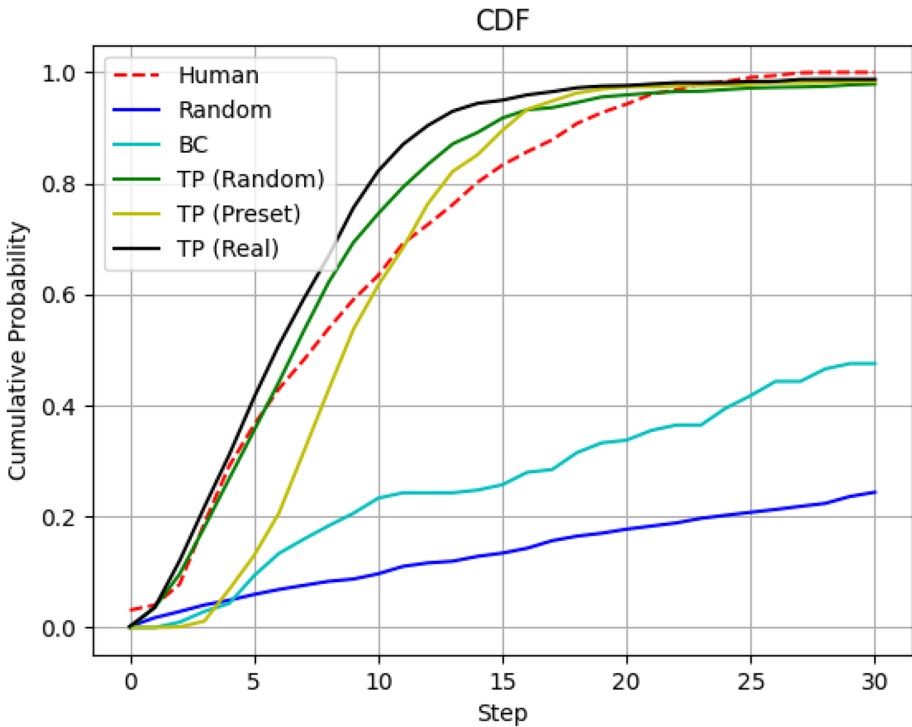


Fig. 4 Cumulative distribution function using logsigmoid activation

method that can be used for comparison, we also trained our generator with the behavioural cloning approach. In Fig. 4 and Table 2, a super-human performance in the efficiency of search can be seen for our method using a logsigmoid activation. Three different initialisation methods, namely Preset, Real and Random, were used. These methods specified which point should be taken as the first point of the trajectory. “Preset” considers a predefined point, such as a point near the checkout, “Real” chooses the first points from real trajectories of the dataset and “Random” picks a random point in the store. All the experiments were performed using an Nvidia GeForce RTX 2080 Ti GPU (11GB of memory) on a 48-CPU Linux machine with Intel(R) Xeon(R) Silver 4214 CPU @ 2.20GHz and 220GB of RAM. The codebase has been developed in Python 3, using the Pytorch library for deep learning. Details about Python requirements are given in the codebase.

Table 2 CDF-AUC results with logsigmoid reward activation

Source	CDF-AUC
Human	21.873
Random	4.139
BC	8.072
TP (Preset)	20.912
TP (Real)	23.735
TP (Random)	22.722

The bold value signifies the best result

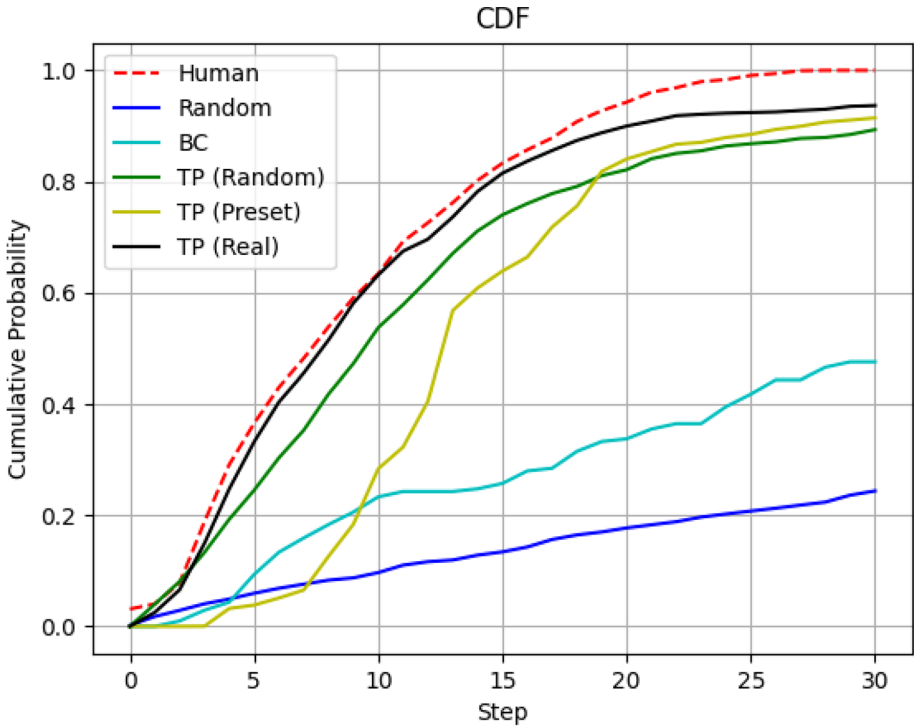


Fig. 5 Cumulative Distribution Function using a linear activation

To overcome the limitations of a non-positive reward, we proposed to use a linear activation. In this, we directly used the output of the discriminator before the sigmoid function to formulate a reward function that had both positive and negative values. The discriminator was trained using the loss proposed by Goodfellow et al. [10], and therefore, its output before the last activation was small and centred on zero. Results of this method are shown in Fig. 5 and Table 3. Our method showed curves that matched more accurately than the human curve, particularly while using the Real initialisation.

The second aspect analysed was the quality of generated trajectories. It was not so trivial to evaluate the quality of a trajectory that was generated completely from scratch without a ground truth. In this work, we proposed to use metrics like Dynamic Time Warping (DTW)

Table 3 CDF-AUC results with linear activation

Origine	CDF-AUC
Human	21.873
Random	4.139
BC	8.072
TP (Preset)	15.999
TP (Real)	20.792
TP (Random)	18.743

The bold value signifies the best result

and Longest Common SubSequence (LCSS) to compare how a certain trajectory is similar to the ones from the test set that have the same task [40].

The DTW distance between two trajectories A and B is calculated according to (7):

$$\begin{aligned} DTW(A, B) = \min(DTW(i-1, j), \\ DTW(i, j-1), \\ DTW(i-1, j-1)) \\ + d(A_i, B_j) \end{aligned} \quad (7)$$

Where: $DTW(i, j)$ is the DTW distance between prefixes of A and B up to positions i and j , A_i and B_j are elements at positions i and j , $d(A_i, B_j)$ is the local distance between A_i and B_j . The LCSS (Longest Common Subsequence) similarity between two sequences A and B is calculated according to (8)[42]:

$$LCSS(A, B) = \frac{|LCS(A, B)|}{\min(|A|, |B|)} \quad (8)$$

Where:

$LCS(A, B)$ is the Longest Common Subsequence of A and B ,
 $|A|$ and $|B|$ are the lengths of sequences A and B .

To calculate an Average DTW distance (ADTW) and an Average LCSS similarity (ALCSS), we chose a generated trajectory and a trajectory from the test set and extracted a subset from each of them that has a starting point near the other (if they exist and if they have more than four points). We repeated this for every trajectory of the test set and every generated trajectory. As a distance function, we used Euclidean distance normalised with two times the diagonal of the store image. Our implementation of LCSS had relaxed constraints: a single point is in common if the Euclidean distance in the discretised grid is less than or equal to $\sqrt{2}$. Then, we normalised the length of the longest common subtrajectory using the length of the generated trajectory. We calculated Similar Trajectories Count (STC) as the number of generated trajectories that match with (distance less than $\sqrt{2}$) at least 50% of the points on a trajectory in the test set. In Table 4, it can be seen that our method had better results than the BC algorithm. The results of the Real and Preset initialisations were also very close to the human results. Human values were taken by splitting the test set into two parts and then comparing one half to the other. It should be noted here that human results were only a reference value and they were not the best result that can be obtained; as these measures evaluate similarity between trajectories, a value higher than the metrics for the human results does not indicate a bad generalisation.

Although our work aims mainly to generate trajectories from zero, we can also use it to forecast existing trajectories. To evaluate the results of this task, we split every test trajectory into two parts and then used the first half to predict the second one. We compared the predicted trajectories with the real ones using Average Displacement Error (ADE) and Final Displacement Error (FDE), two widely used metrics for forecasting trajectories [18].

ADE is a common metric used to evaluate the accuracy of trajectory predictions in the field of computer vision and robotics. It measures the average Euclidean distance between the predicted positions and the ground truth positions of objects over a sequence of time steps. ADE is calculated as shown in (9):

$$ADE = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \quad (9)$$

Table 4 Results of similarity measures

Results			
Source	ADTW ↓	ALCSS ↑	STC ↑
Human	0.478	0.463	0.114
Random	2.274	0.122	0.005
BC	0.782	0.304	0.072
TP (Preset)	0.611	0.441	0.102
TP (Real)	0.527	0.417	0.094
TP (Random)	0.477	0.429	0.023

The bold values signify the best results

where N is the number of time steps, (x_i, y_i) are the ground truth positions, and (\hat{x}_i, \hat{y}_i) are the predicted positions at time step i . A lower value means that the predicted positions are closer to the ground truth positions on average over the entire trajectory. FDE is another important metric used to evaluate trajectory predictions, particularly at the final time step. It measures the Euclidean distance between the predicted final position and the ground truth final position of an object. FDE is calculated as shown in (10):

$$FDE = \sqrt{(x_N - \hat{x}_N)^2 + (y_N - \hat{y}_N)^2} \quad (10)$$

where N is the final time step, (x_N, y_N) is the ground truth final position, and (\hat{x}_N, \hat{y}_N) is the predicted final position. For ADE, a lower value means that the predicted positions are closer to the ground truth positions on average over the entire trajectory. In other words, a lower ADE indicates that the trajectory predictions are more accurate. For FDE, a lower value means that the predicted final position is closer to the ground truth final position. This is particularly important when evaluating the accuracy of predictions at the end of a trajectory. A lower FDE indicates better accuracy in predicting the final destination of an object.

In Table 5, we can see how our framework obtained better results than the behaviour cloning method. Moreover, it differed considerably from the results of the random prediction. However, metrics like ADE and FDE do not consider multimodality, as they compare a forecast with only one of the possible real trajectories. So, these metrics should be used only for a qualitative comparison. Qualitative results on prediction and forecasting are available in the appendix.

5 Limitations

The limitations of the proposed approach primarily revolve around its generalization capabilities concerning stores with significant differences in layout and size. The need for a

Table 5 Forecasting results

Results		
Source	ADE	FDE
Random	0.133	0.141
BC	0.058	0.053
TP	0.033	0.035

The bold values signify the best results

predefined fixed number of cells for the map discretization can have varying impacts on the categorization of items, depending on the layout of each specific store. Another limitation is that the collected dataset takes into consideration only shoppers with either a shopping cart or a trolley, which can exhibit considerably different behaviour from a shopper without.

6 Conclusions

In this paper, it is proposed GREEN PATH, an intelligent expert system for space planning that employs a GAIL framework for modelling human trajectories in an environment. This work allowed for both generating trajectories from scratch and predicting the future patterns of a person from existing trajectories. The system is a predictive and generative model that can handle an environment with complex constraints, such as those in retail. In particular, a GAIL-based framework has been hybridised with classical reinforcement learning methods, such as continuous penalties, which allow for modelling the shape of the trajectories and inserting a bias in the training. The system is also very general, and the data can be constructed in multiple ways. Depending on the chosen reward, we can either enhance or emulate the behaviour of a human being. This paper focused on the second aspect, as it was more interesting for our purposes. Finally, the experimental results clearly show the feasibility of the proposed method as well as its generalisability, since state is based on the exploration, it is possible to generalize its creation during the deployment phase and also carry it out through other methods, such as visual input from a robot. Therefore, with such a framework, it is possible to develop a store simulator where we can predict customer behaviour with different layouts and shelf positions.

7 Future works

Future works will be devoted to improving the results based on the limitations highlighted in Section 5. A higher number of stores in the dataset will surely be mandatory for this task. However, if the analysis focused on a single store, the framework could provide better results if it were trained with fewer stores with many trajectories. This topic should be subject to further analysis. On the framework side, the implementation of different state and action spaces that could provide better results should be taken into consideration. A new paradigm of space design will be achieved. There will be an increased number of public spaces that will re-arrange their layout following the data collected by GREEN PATH. Managers will increase their knowledge of space utilization, data will be shared at a worldwide scale to define a shared protocol among designers. These technologies could also be integrated into a user-intuitive framework, designed around the challenges previously described, ultimately enabling a system that can be integrated seamlessly into existing spaces, even for other domains, without the need to fully re-engineer the existing environments for visitors.

Appendix

A - Generated trajectories

The following Figures aim to describe the qualitative results of generated trajectories. They refer to the main categories of the store's target. The dataset is collected in different important supermarkets. In particular, Fig. 6 depicts an example of a generated trajectory with a task



Fig. 6 Generated trajectory with “meat” task and “random” initialization

that considers the meat category with random initialization. The trajectory starts at a random point near the cashout, passes through the store and then arrives in the task zone passing around the cold department.

Figure 7 represents another example. The category chosen for this task is beer with real initialization. In this case, the actor has not found the task zone in the maximum time specified. It changed its mind many times and explored various categories without finding the beer, that is located in the left-bottom zone.



Fig. 7 Generated trajectory with “beer” task and “real” initialization

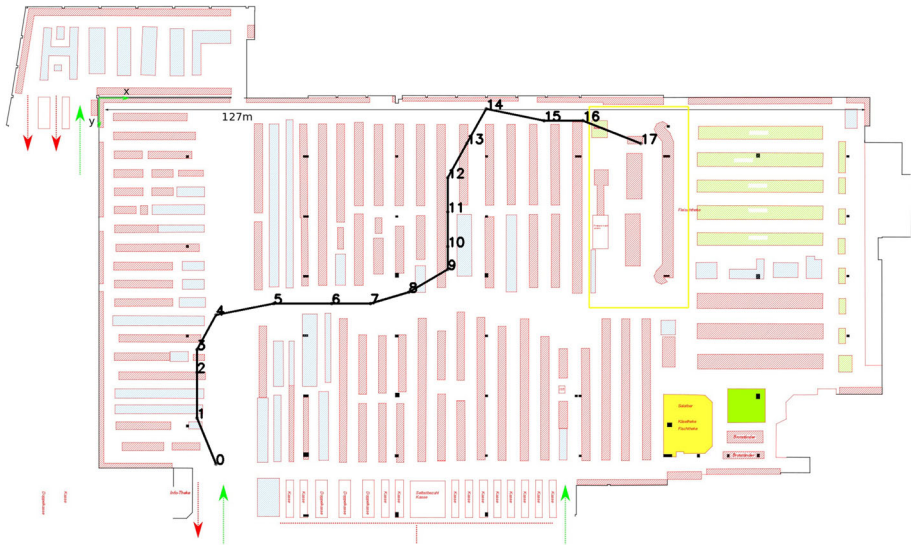


Fig. 12 Generated trajectory with “grocery” task and “first” initialization

Another important category in the stores in which our experiments have been assessed is fish. Thus, Fig. 8 shows the generated trajectory with “fish” task and “real” initialization. In this example, it is shown how the constraints for walls are soft. The agent crosses a wall to reach the task area. This is not a problem in our case, because what we want is to analyze behaviours in a certain context, rather than a realistic one. However, these constraints can be easily hardened in post-processing. Figure 9, instead, depicts the generated trajectory with “fish” task and “first” initialization. Hence, the actor shows a relatively complex behaviour changing its main direction several times. Figure 10 is the last example for this category that shows the generated trajectory with “fish” task and “random” initialization. The generator chooses a relatively longer path instead of the trivial one.

Following, the category devoted to breakfast products, Fig. 11 aims at representing the generated trajectory with the breakfast task and first initialization. The actor changed his mind in the first steps and then he arrived easily at the task zone.

Finally, Fig. 12 reports the generated trajectory with “grocery” task and “first” initialization. Many customers choose to go through the central corridor, the actor correctly modelled such behaviour.

B - Forecasted trajectories

In this section, qualitative results for forecasted trajectories are reported for a complete analysis of the proposed approach. The evaluation has been performed for the most significative categories in the store taken into exam. Figure 13 is an example of a forecasted trajectory with “beer” task.

Figure 14 is the one devoted to the breakfast category. In particular, it depicts the forecasted trajectory with the “breakfast” task. In this case, forecasting the trajectory is not similar. However, it is not fair to say that this forecasting is “wrong”. It clearly shows another behaviour that can be real, taking into account multimodality.

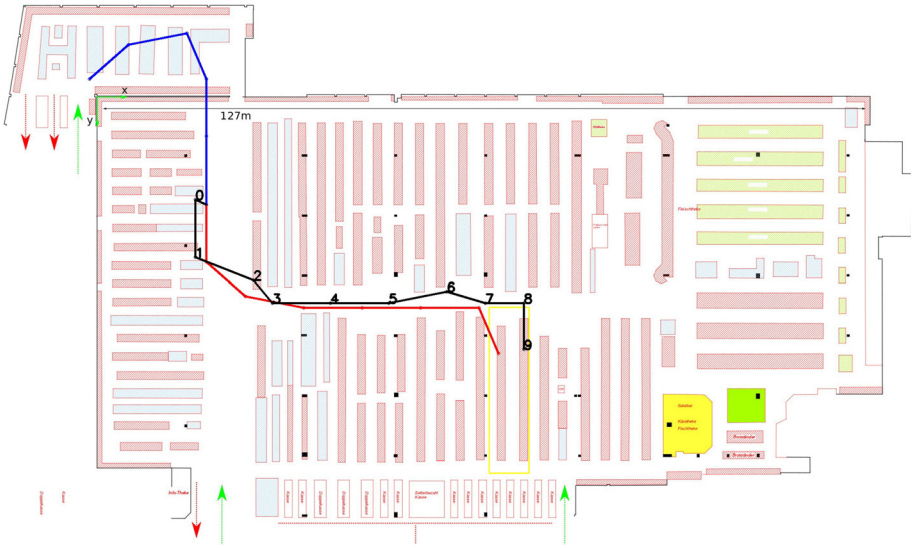


Fig. 13 An example of a forecasted trajectory with “beer” task



Fig. 14 Forecasted trajectory with “breakfast” task

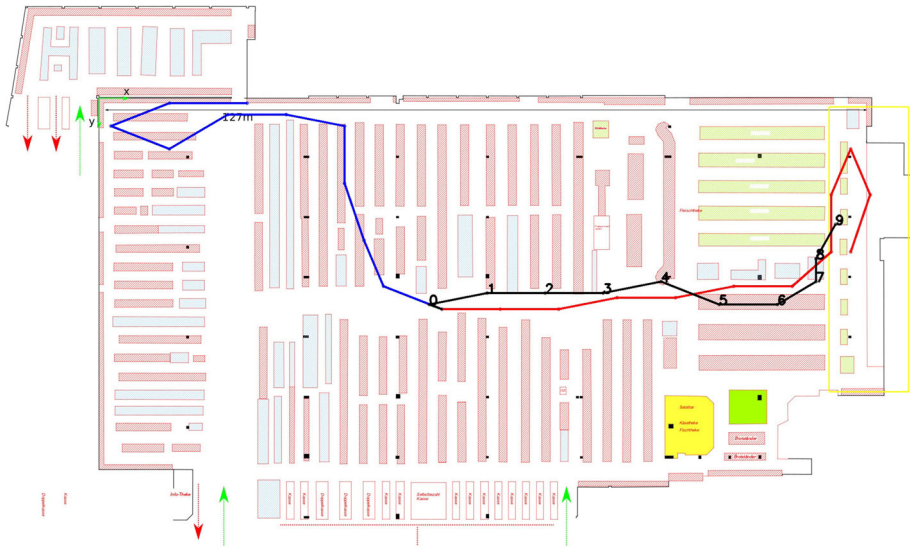


Fig. 15 A successfully forecasted trajectory with “meat” task

Then, Fig. 15 shows a successfully forecasted trajectory with the “meat” task.

Acknowledgements This work was funded by Grottini Lab (www.grottinilab.com).

Funding Open access funding provided by Università Politecnica delle Marche within the CRUI-CARE Agreement.

Data Availability Data and codebase used in this study available at <https://github.com/rokopi-byte/greenpath>.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alahi A, Goel K, Ramanathan V, Robicquet A, Fei-Fei L, Savarese S (2016) Social Istm: Human trajectory prediction in crowded spaces. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 961–971
2. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In: International conference on machine learning. pp. 214–223, PMLR

3. Ballan L, Castaldo F, Alahi A, Palmieri F, Savarese S (2016) Knowledge transfer for scene-specific motion prediction. In: European Conference on Computer Vision. Springer, pp. 697–713
4. Bartoli F, Lisanti G, Ballan L, Del Bimbo A (2018) Context-aware trajectory prediction. In: 2018 24th International Conference on Pattern Recognition (ICPR). pp. 1941–1946, IEEE
5. Demetriou A, Alfsvåg H, Rahrovani S, Haghiri Chehreghani M (2023) A deep learning framework for generation and analysis of driving scenario trajectories. *SN Comput Sci* 4(3):251
6. Englert P, Vien NA, Toussaint M (2017) Inverse kkt: Learning cost functions of manipulation tasks from demonstrations. *The Int J Robot Res* 36(13–14):1474–1488
7. Ferracuti N, Norscini C, Frontoni E, Gabellini P, Paolanti M, Placidi V (2019) A business application of rtls technology in intelligent retail environment: Defining the shopper's preferred path and its segmentation. *J Retail Consum Serv* 47:184–194
8. Gabellini P, D'Aloisio M, Fabiani M, Placidi V (2019) A large scale trajectory dataset for shopper behaviour understanding. In: International conference on image analysis and processing. Springer, pp. 285–295
9. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, pp. 249–256
10. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. *Commun ACM* 63(11):139–144
11. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC (2017) Improved training of wasserstein gans. *Adv Neural Inf Process Syst* 30
12. Gupta A, Johnson J, Fei-Fei L, Savarese S, Alahi A (2018) Social gan: Socially acceptable trajectories with generative adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2255–2264
13. Ho J, Ermon S (2016) Generative adversarial imitation learning. *Adv Neural Inf Process Syst* 29
14. Huang S, Li X, Zhang Z, He Z, Wu F, Liu W, Tang J, Zhuang Y (2016) Deep learning driven visual path prediction from a single image. *IEEE Trans Image Process* 25(12):5892–5904
15. Hussein A, Gaber MM, Elyan E, Jayne C (2017) Imitation learning: A survey of learning methods. *ACM Comput Surv (CSUR)* 50(2):1–35
16. Karras T, Laine S, Aila T (2021) A style-based generator architecture for generative adversarial networks. *IEEE Trans Pattern Anal Mach Intell* 43(12):4217–4228
17. Kim B, Pineau J (2016) Socially adaptive path planning in human environments using inverse reinforcement learning. *Int J Soc Robot* 8(1):51–66
18. Korbmacher R, Tordeux A (2022) Review of pedestrian trajectory prediction methods: Comparing deep learning and knowledge-based approaches. *IEEE Trans Intell Transp Syst*
19. Kosaraju V, Sadeghian A, Martín-Martín R, Reid I, Rezatofighi H, Savarese S (2019) Social-bigat: Multimodal trajectory forecasting using bicycle-gan and graph attention networks. *Adv Neural Inf Process Syst* 32
20. Kothari P, Kreiss S, Alahi A (2021) Human trajectory forecasting in crowds: A deep learning perspective. *IEEE Trans Intell Transp Syst*
21. Kuettel D, Breitenstein MD, Van Gool L, Ferrari V (2010) What's going on? discovering spatio-temporal dependencies in dynamic scenes. In: 2010 IEEE computer society conference on computer vision and pattern recognition. pp. 1951–1958, IEEE
22. Li H, Jiao H, Yang Z (2023) Ship trajectory prediction based on machine learning and deep learning: A systematic review and methods analysis. *Eng Appl Artif Intell* 126:107062
23. Li Y, Song J, Ermon S (2017) Infogail: Interpretable imitation learning from visual demonstrations. *Adv Neural Inf Process Syst* 30
24. Ma WC, Huang DA, Lee N, Kitani KM (2017) Forecasting interactive dynamics of pedestrians with fictitious play. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 774–782
25. Mescheder L, Geiger A, Nowozin S (2018) Which training methods for gans do actually converge? In: International conference on machine learning. pp. 3481–3490, PMLR
26. Morrissey J, Moore T, Horne RE (2011) Affordable passive solar design in a temperate climate: An experiment in residential building orientation. *Renew Energy* 36(2):568–577
27. Ng AY, Harada D, Russell S (1999) Policy invariance under reward transformations: Theory and application to reward shaping. In: *Icml*, vol. 99. Citeseer, pp. 278–287
28. Paolanti M, Liciotti D, Pietrini R, Mancini A, Frontoni E (2018) Modelling and forecasting customer navigation in intelligent retail environments. *J Intell Robot Syst* 91(2):165–180

29. Paolanti M, Pietrini R, Mancini A, Frontoni E, Zingaretti P (2020) Deep understanding of shopper behaviours and interactions using rgb-d vision. *Mach Vision Appl* 31(7):66. <https://doi.org/10.1007/s00138-020-01118-w>
30. Pei Z, Qi X, Zhang Y, Ma M, Yang YH (2019) Human trajectory prediction in crowded scene using social-affinity long short-term memory. *Pattern Recognit* 93:273–282
31. Pellegrini S, Ess A, Schindler K, Van Gool L (2009) You'll never walk alone: Modeling social behavior for multi-target tracking. In: 2009 IEEE 12th international conference on computer vision. pp. 261–268, IEEE
32. Pierdicca R, Paolanti M, Vaira R, Marcheggiani E, Malinverni ES (2019) Frontoni E (2019) Identifying the use of a park based on clusters of visitors' movements from mobile phone data. *J Spatial Inf Sci* 19:29–52
33. Robicquet A, Sadeghian A, Alahi A, Savarese S (2016) Learning social etiquette: Human trajectory understanding in crowded scenes. In: European conference on computer vision. Springer, pp. 549–565
34. Rossi L, Paolanti M, Pierdicca R, Frontoni E (2021) Human trajectory prediction and generation using lstm models and gans. *Pattern Recognit* 108136
35. Roth K, Lucchi A, Nowozin S, Hofmann T (2017) Stabilizing training of generative adversarial networks through regularization. *Adv Neural Inf Process Syst* 30
36. Schulman J, Moritz P, Levine S, Jordan M, Abbeel P (2018) High-dimensional continuous control using generalized advantage estimation
37. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O (2017) Proximal policy optimization algorithms
38. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M et al (2016) Mastering the game of go with deep neural networks and tree search. *Nat* 529(7587):484–489
39. Solera F, Calderara S, Cucchiara R (2015) Learning to divide and conquer for online multi-target tracking. In: Proceedings of the IEEE international conference on computer vision. pp. 4373–4381
40. Tao Y, Both A, Silveira RI, Buchin K, Sijben S, Purves RS, Laube P, Peng D, Toohey K, Duckham M (2021) A comparative analysis of trajectory similarity measures. *GISci Remote Sensing* 58(5):643–669
41. Tuan YF (1979) Space and place: humanistic perspective. In: *Philosophy in geography*. Springer, pp. 387–427
42. Vlachos M, Kollios G, Gunopulos D (2002) Discovering similar multidimensional trajectories. In: Proceedings 18th international conference on data engineering. pp. 673–684, IEEE
43. Walker J, Doersch C, Gupta A, Hebert M (2016) An uncertain future: Forecasting from static images using variational autoencoders. In: European conference on computer vision. Springer, pp. 835–851
44. Walker J, Gupta A, Hebert M (2014) Patch to the future: Unsupervised visual prediction. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3302–3309
45. Wu S, Yang H, Zheng S, Su H, Fan Y, Yang MH (2017) Crowd behavior analysis via curl and divergence of motion trajectories. *Int J Comput Vision* 123(3):499–519
46. Yang Z, Huang L, Chen Y, Wei Z, Ahn S, Zelinsky G, Samaras D, Hoai M (2020) Predicting goal-directed human attention using inverse reinforcement learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 193–202
47. Zamboni S, Kefato ZT, Girdzijauskas S, Norén C, Dal Col L (2022) Pedestrian trajectory prediction with convolutional neural networks. *Pattern Recognit* 121:108252

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.