IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 March 2024, accepted 3 April 2024, date of publication 12 April 2024, date of current version 19 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3387698

RESEARCH ARTICLE

POI-GAN: A Pedestrian Trajectory Prediction Method for Service Scenarios

YE LI, CHI ZHANG, JINGKANG ZHOU, AND SHENGCUI ZHOU
University of Shanghai for Science and Technology, Shanghai 200093, China

Corresponding author: Ye Li (liye@usst.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61703277, and in part by Shanghai Sailing Program under Grant 17YF1427000.

ABSTRACT In the service industry, some operations require preparatory measures and cannot deliver immediate and reliable services. Identifying potential clients in densely populated environments presents a significant challenge for improving service efficiency. Current methodologies used to predict pedestrian trajectories demonstrate suboptimal performance in service-oriented contexts. In response to these challenges, this paper introduces ''POI-GAN'', a novel approach tailored to forecasting pedestrian trajectories in service-centric settings. POI-GAN devises an interest point model that conceptualizes service locations within the scene as distinct obstacles. These service locations are subsequently characterized employing a social force model. Additionally, the framework introduces a field of view angle model which filters interactions between dynamic and static objects in the scene to establish plausibility. Subsequently, a generative model is used to produce projected pedestrian trajectories for future time frames. Empirical evaluations highlight the effectiveness of POI-GAN in improving trajectory prediction, particularly in scenarios with multiple interest points in the scene. Notably, POI-GAN exhibits superior performance when measured against analogous methods, as evidenced by the improved Average Displacement Error (ADE) and Final Displacement Error (FDE) metrics. This innovative approach empowers service providers with the capacity to effectively discern potential customers within the scene, ultimately elevating the quality of service delivery.

INDEX TERMS Pedestrian trajectory prediction, generative adversarial networks, deep learning.

I. INTRODUCTION

In recent years, deep learning-based pedestrian trajectory prediction has emerged as a highly researched topic in the fields of artificial intelligence and computer vision. This technology holds significant potential due to its wide-ranging applications in intelligent transportation, smart security, and other domains. In service-oriented scenarios, certain services require advance preparation, and failure to complete this preparation before the client's arrival can result in wait times, where the average wait time serving as a key metric for assessing service quality [1]. Furthermore, proactive identification of potential clients within a given environment, coupled with tailored service provision, contributes to enhancing service efficiency [2]. Goal-oriented pedestrian trajectory prediction algorithms enable models to identify potential clients within a setting, empowering them to comprehend the behavioral intentions represented by partial trajectories and to prepare adequately before the client arrives at the service point [3].

For some automated services, optimization through pedestrian trajectory prediction enhances service quality and saves costs for service providers. For instance, by predicting user intentions, vehicles within parking facilities can initiate engine startup, adjust internal temperatures, and process parking fees in advance. Brand marketers can leverage pedestrian trajectory prediction to deliver targeted advertising, increase the frequency of core content playback, and offer more detailed content recommendations upon detecting potential target clients. In virtual reality (VR) and augmented

The associate editor coordinating the review of this manuscript and approving it for publication was Kathiravan Srinivasan.

reality (AR) applications, accurately predicting user focal points and prioritizing the allocation of computational resources improve user experience while simultaneously reducing computational costs.

Pedestrian trajectory prediction entails inferring the future position coordinates for a target pedestrian, which is achieved through an analysis of their historical trajectory [4] and behavioral characteristics [5]. However, the complexity of pedestrian movement arises not only from environmental influences but also individual social habits [6], making pedestrian trajectory prediction a formidable challenge.

The widespread application of social force models enables current mainstream pedestrian trajectory prediction algorithms to understand social behaviors between pedestrians and between pedestrians and objects [3]. However, these objects are commonly classified as general obstacles and pathways, encompassing entities such as vehicles, streetlights, lawns, and sidewalks. Pedestrians tend to exhibit avoidance behavior when encountering these general obstacles, as illustrated in Figure 1(a). Conversely, in service scenarios, pedestrians may be drawn to specific service points and actively approach them, exemplified by billboards, park benches, and snack stalls, as depicted in Figure 1(b). However, from a physical perspective, these service points may also be considered as obstacles. Using the social modeling framework designed for general obstacles to represent them results in predicted pedestrian trajectories that diverge from reality. Specifically, these trajectories tend to avoid these service points, which contradicts the actual behavior of pedestrians actively seeking services, as shown in Figure 1(c). Existing research often fails to adequately address the intricate interaction dynamics between pedestrians and service points. Moreover, mainstream datasets commonly used in trajectory prediction, such as ETH and UCY, primarily focus on collecting data at densely populated intersections, neglecting scenarios where pedestrians pause to receive services. Therefore, to investigate the impact of service points on pedestrian trajectories, it is imperative to collect trajectory information in specific scenes.

A major challenge in pedestrian trajectory prediction lies in the model's potential limitation in perceiving data dimensions. Especially when service points with low interaction volumes are present, the model's capacity to learn pedestrian trajectories around these points diminishes. Pedestrian trajectories that interact with these service points hold considerably higher value than those that simply pass by without interaction. Improving the performance of pedestrian trajectory prediction near service points requires exposing the model to a greater number of these high-value trajectories.

To investigate the impact of service points on pedestrian trajectory prediction within a scene, we propose a novel method, POI-GAN, which integrates the interest point model, field of view angle model, and generative adversarial network. This method utilizes historical pedestrian trajectories as input to forecast pedestrian movement in the upcoming seconds.

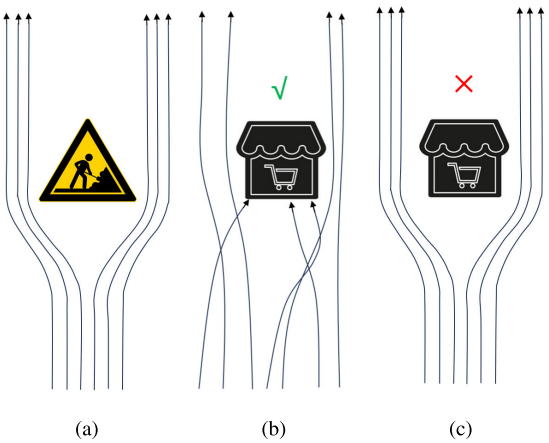


FIGURE 1. The impact of general obstacles and service points on pedestrian trajectories.

The primary contributions of this research are as follows:

(1) The abstraction of service points within the scene as a distinct category of obstacles, termed ''Points of Interest'' (POI). In contrast to traditional models that treat all obstacles equally, POIs can attract nearby pedestrians and possess attributes typical of obstacles, such as collision volumes and pedestrian impassability. The utilization of the social force model enables the modeling of POIs, facilitating their rational participation in social computations.

(2) The establishment of a field of view model to filter interactions among interest points and pedestrians, as well as among pedestrians themselves. This model mitigates unrealistic interactions prior to the social pooling operation, thereby enhancing trajectory plausibility.

(3) The development of the Point of Interest - Generative Adversarial Network (POI-GAN), which integrates the proposed POI model and field of view angle model into the Info-GAN [16] framework. Through continuous input of generated pedestrian trajectories alongside real trajectory information into the discriminator network, the generator and discriminator engage in adversarial training to iteratively optimize both network parameters, thereby enabling the generator to produce high-quality predicted trajectories.

(4) The adoption of importance sampling algorithms during the training phase to increase the likelihood of selecting high-value samples in the initial stages of training. This measure reduces the required training epochs and accelerates model convergence.

(5) The proposal of a data collection scheme based on millimeter-wave radar and the HomeAssistant platform. Compared to traditional approaches utilizing cameras for data collection, the new scheme eliminates the need for pedestrian trajectory recognition in images, significantly reducing the investment costs of perception hardware and computational hardware. Additionally, manual steps such as image annotation in dataset

creation are eliminated, thereby improving dataset creation efficiency.

Empirical analysis reveals that the POI-GAN model substantially enhances the accuracy and interpretability of pedestrian trajectory prediction within service scenarios. In test environments featuring interest points, the POI-GAN model outperforms other comparable methods [14], [15], [37] with respect to ADE and FDE indicators. Moreover, in another indoor automated service scenario, POI-GAN exhibited robust pedestrian intention discrimination capabilities, leading to improved service quality in automated systems.

II. RELATED WORKS

Current pedestrian trajectory prediction approaches can be mainly categorized into two categories: shallow learning-based prediction methods and deep learning-based prediction methods. Some review articles [3], [7], [8], [9], [10] provide detailed introductions to classical methods in this field as well as the latest research endeavors. A seminal contribution in this domain is the Social Force model, proposed by Helbing and Molnar [11], which characterizes pedestrian motion trajectories by considering interactions between pedestrians and environmental factors, with social forces being a primary determinant of pedestrian behavior. Alahi et al. [12] proposed a method to capture social affinity characteristics among pedestrians by learning their relative positions within crowds. This approach leverages changes in relative positions to describe pedestrian interactions, thereby enhancing trajectory prediction accuracy. Yi et al. [13] introduced a method that enhances crowd behavior prediction by integrating human attributes as inputs. By using human attributes for prediction, this method offers improved descriptions of behavioral characteristics within various crowd contexts, consequently enhancing prediction accuracy. However, prediction methods based on shallow learning have limitations in representing the complete spectrum of complex and abstract interactions among pedestrians. Consequently, these methods often fall short in capturing the intricacies of real-world scenarios, making them less effective in predicting complex pedestrian motion models and scenarios.

The proliferation of deep learning, driven by advancements in hardware computing power, has led to the dominance of prediction methods based on deep learning that utilize historical data. Deep neural networks inherently have the capability to autonomously learn pedestrian interactions, eliminating the necessity for predefined pedestrian interaction behaviors—a limitation of shallow learning methods. Alahi et al. [14] introduced the Social LSTM pedestrian trajectory prediction method, which models each pedestrian's historical trajectory as an independent LSTM layer. Adjacent pedestrians can share hidden states through a social pooling layer, facilitating interaction modeling among pedestrians. Amirian et al. [15] proposed the Social Ways generative adversarial model, inspired by Info-GAN [16], which not only addressed GAN training challenges but also incorporated an attention mechanism to enable autonomous allocation of attention to interaction information, further enhancing trajectory prediction accuracy. Liu et al. [17] introduced CoL-GAN, which integrated an attention mechanism into the generative adversarial network and employed a convolutional neural network as a discriminator, enabling the model to generate collision-free trajectories. Huang et al. [18] devised STI-GAN, which harnessed generative adversarial networks and a graph attention network (GAT) based on pedestrian spatio-temporal interaction information. This approach simulated pedestrian distribution to capture prediction path uncertainty and generate more plausible future trajectories. Mohamed et al. [19] introduced Social-STGCNN, adopting a distinct research approach by replacing the original model's recurrent recursive architecture with a graph convolutional network. Chiho et al. [20] proposed a goal-oriented approach named DROGON, which utilizes trajectory data to achieve higher-level objectives, specifically the inference of driver and pedestrian intentions. In recent years, with the optimization of computing hardware, methods based on graph convolutional networks [21], [22], [23], [24], [25], [26], [27] have shown significantly faster inference speed compared to time series models [14], [28] and generative models [15], [16], [17], [18], [29].

Insufficient data volume is a notable hindrance that limits improvements in pedestrian trajectory prediction accuracy. Consequently, some researchers have augmented prediction by incorporating additional data dimensions, such as motion posture [30], [31], [32], [33], semantic gaps [34], and other information to aid inference.

III. PEDESTRIAN TRAJECTORY PREDICTION ALGORITHM BASED ON POINT OF INTEREST MODEL

A. PEDESTRIAN TRAJECTORY PREDICTION PROBLEM DEFINITION

Pedestrians typically adhere to pre-planned routes in unobstructed environments, but on complex roads, they are often influenced by other pedestrians and environmental factors. However, on complex traffic roads, they are often influenced by pedestrians and environmental factors. Pedestrians consider various factors in their own path planning and may modify their intended route based on the surrounding environment, as well as the direction and speed of other pedestrians and vehicles.

Pedestrian trajectory prediction problem can be regarded as a problem of predicting the future path based on the current pedestrian's historical path and historical information, which is essentially a prediction problem based on time series [7]. We defines the pedestrian's trajectory $X = (X_1, X_2, \ldots, X_n)$, where $X_i = \left\{\left(x_i^t, y_i^t\right) \mid t \epsilon (1, \cdots, t_{obs})\right\}$, t is the trajectory frame, t_{obs} is the first frame at the beginning of the observation, n is the number of all target pedestrians in the observed scene, $\left(x_i^t, y_i^t\right)$ is the two-dimensional plane coordinate of the target pedestrian at time, $\left(x_i^t, y_i^t\right)$ is the observation time sequence length, and the pedestrian's real trajectory can be expressed as:

$$Y = (Y_1, Y_2, \ldots, Y_n) \tag{1}$$

$$Y_i = \left\{(x_i^t, y_i^t)|t \in (t_{obs}+1, t_{obs}+2, \ldots, t_{obs}+t_{pred})\right\} \quad (2)$$

where t_{pred} is the prediction time sequence length. Similarly, the pedestrian trajectory predicted can be expressed as:

$$\hat{Y} = \left(\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n\right) \quad (3)$$

$$\widehat{Y}_i = \{(x_i^t, y_i^t)|t \in \left(t_{obs}+1, t_{obs}+2, \ldots, t_{obs}+t_{pred}\right)\} \quad (4)$$

B. MODEL STRUCTURE

POI-GAN is based on Social-ways trajectory prediction algorithm, which itself is built upon the framework of Info-GAN [16]. Its main components consist of a trajectory generator (Generator) and a trajectory discriminator (Discriminator). Within the generator, POI-GAN incorporates modules optimized for service scenarios, namely the Point of Interest module (POI Module) and the Field of View module (FOV Module). Figure 2 depicts the overall algorithm framework. Info-GAN tackles the issue of controlling the GAN's generation distribution by adjusting latent code tendencies in unsupervised settings. Unlike traditional GANs, it concentrates on unsupervised learning without data labels, introducing two novel components into the GAN network: latent code and information loss. Throughout training, the discriminator evaluates the authenticity of the generated data, calculates errors through a loss function, and provides feedback to different segments of the model to adjust parameters. After training, the discriminator becomes inactive, and the generator directly produces predicted trajectories using input pedestrian historical trajectory data.

POI-GAN employs LSTM networks [14] to capture the complex dynamics of pedestrian movements and interactions in crowded environments. Both the generative and discriminative models integrate LSTM networks to fulfill their respective objectives.

The generator accepts input data consisting of Gaussian noise Z, latent code C, and pedestrian trajectory data samples from the dataset. Latent code C, a new input compared to traditional GANs, standardizes the output. The generator's core employs an Encoder-Decoder architecture. G-Encoder integrates multiple LSTM-PED pedestrian historical trajectory feature encodings. The feature vector for pedestrian i undergoes encoding via the trajectory encoder, employing a fully connected neural network layer with ReLU nonlinear activation. Apart from pedestrians, the interest point module (POI Module) abstracts predefined interest points in the scene as LSTM-POI encodings. LSTM-PED encoding and LSTM-POI encoding interact via the field of view angle module (FOV Module) to screen for interaction behaviors. This step eliminates unreasonable interactions before entering the social pooling module for social computation, providing interaction information. Noise Z, latent code C, pedestrian trajectory encoding LSTM-PED, and interaction information serve as the four inputs for G-Decoder, which decodes and outputs the pedestrian trajectory prediction coordinate sequence $\widehat{Y}_{i:T}$.

The discriminator discerns between true and false trajectories. It simultaneously receives the predicted trajectories from the trajectory generator and the actual pedestrian trajectories concurrently, producing a high-dimensional feature vector via a fully connected neural network layer (FC). Subsequently, LSTM networks and a multilayer perceptron calculate the probability of the generated trajectory's authenticity. The output of the fully connected layer is divided into D and Q components. Info-GAN diverges from the $L2$ loss function in Social GAN by incorporating a subnetwork Q within the discriminator to generate information loss. Q functions as a latent code retriever, where information loss is defined as the Mean Squared Error (MSE) between the recovered latent code $\hat{c}$ and the actual latent code c.

The loss function consists of two components: $L_{GAN}(G, D)$ and $L_{L_1}(G, Q)$. $L_{GAN}(G, D)$ represents the generative adversarial network's loss function, encompassing discrimination and adversarial losses. $L_{L_1}(G, Q)$ denotes the mutual information between the latent code and generated samples, essentially indicating information loss.

C. POINT OF INTEREST MODEL

Amidst complex traffic environments, pedestrians often consider the speed, direction, and relative Euclidean distance of nearby individuals when determining their movement trajectory. Consequently, they adapt their trajectory based on their own judgment. Simultaneously, environmental factors and changes in neighboring pedestrians' trajectories can influence other pedestrians. Consequently, pedestrians' trajectories evolve continually over time due to these influences. In service scenarios, pedestrian movement trajectories are influenced not only by fellow pedestrians, roads, and conventional obstacles but also by specific service points.

To simulate the influence of pedestrian interactions with service points on their movement trajectories, we conceptualize service points as interest points using the social force model [11]. The force exerted by interest point i on surrounding pedestrians can be described as follows:

$$\vec{F}_\alpha(t) = \overrightarrow{f_{\alpha_\beta}}\left(\vec{r}_{\alpha\beta}\right) + \overrightarrow{f_{\alpha_i}}\left(\left\|\overrightarrow{r_{\alpha_i}}\right\|, t\right) \quad (5)$$

$$\overrightarrow{f_{\alpha\beta}}\left(\vec{r}_{\alpha\beta}\right) = -\Delta\overrightarrow{r_\alpha}_\beta V_{\alpha\beta}\left[b\left(\vec{r}_{\alpha\beta}\right)\right] \quad (6)$$

$$\overrightarrow{f_{\alpha_i}}\left(\left\|\overrightarrow{r_{\alpha_i}}\right\|, t\right) = -\Delta\overrightarrow{r_\alpha}_i W_{\alpha i}\left(\|\vec{r}_{\alpha i}\|, t\right) \quad (7)$$

where:

$\overrightarrow{f_{\alpha_i}}\left(\left\|\overrightarrow{r_{\alpha_i}}\right\|, t\right)$ represents the attraction force exerted by interest point i on the surrounding pedestrians at time t, with r being the action radius. Pedestrians within a certain range are drawn toward interest points.

$\overrightarrow{f_{\alpha\beta}}\left(\vec{r}_{\alpha\beta}\right)$ represents the repulsion force exerted by interest point i on surrounding pedestrians at time t, with r being the action radius. Pedestrians within an r_2 range are repelled by the obstacle property inherent to interest points.

The interaction dynamics between interest points and pedestrians can be classified into three scenarios, illustrated in Figure 3:

(1) If the Euclidean distance $r_t^{i,j}$ between pedestrian j and interest point i exceeds r_l, the movement trajectory of this pedestrian remains unaffected by interest point p_l.

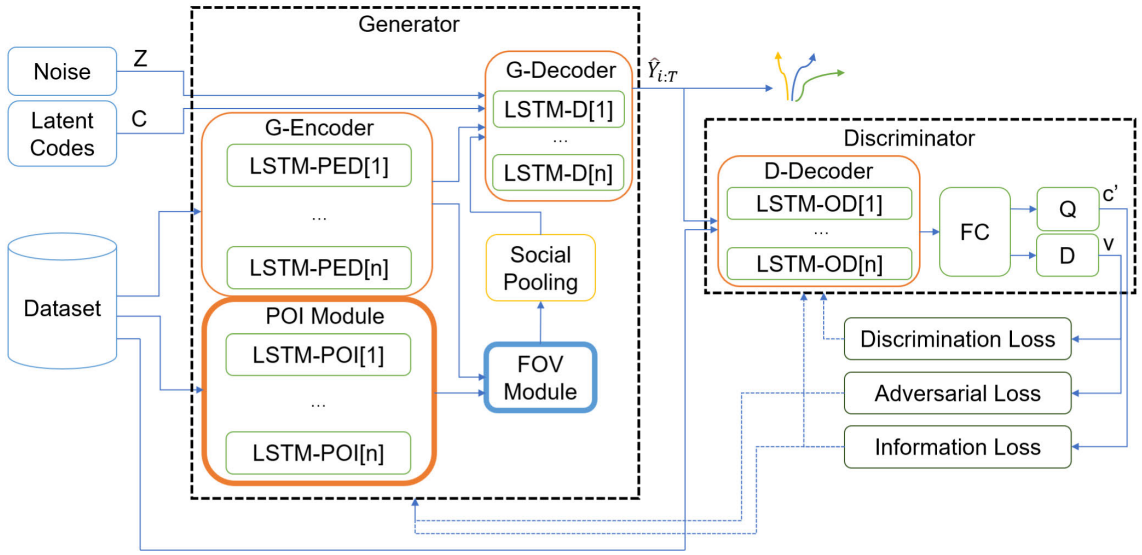


FIGURE 2. POI-GAN framework.

(2) If the Euclidean distance between pedestrian and interest point is greater than r_l but less than r_2, the pedestrian will be attracted by interest point p_1. If the Euclidean distance between a pedestrian and an interest point is less than r_2, the pedestrian will be impeded by the obstacle posed by the interest point, making further approach difficult.

(3) If the Euclidean distance between a pedestrian and an interest point is less than r_2, the pedestrian encounters obstacles presented by the interest point, making further approach difficult.

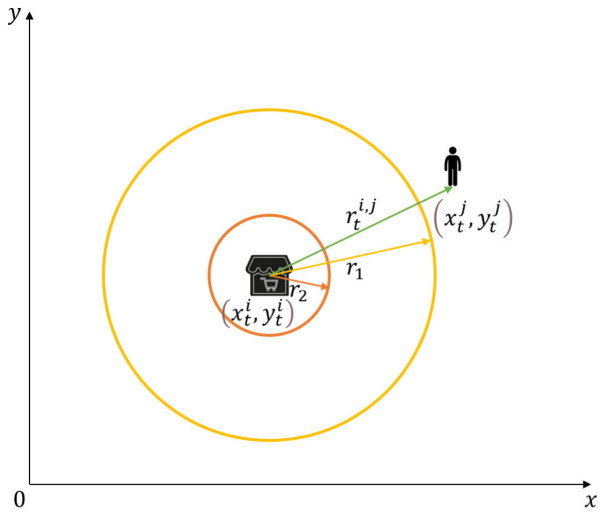


FIGURE 3. Interaction between points of interest and pedestrians.

D. FIELD OF VIEW MODEL

During interactions with interest points and pedestrians, as well as interactions among pedestrians themselves, individuals tend to direct their focus towards those directly in front of them. They glean pertinent information within their field of vision by adjusting their head orientation. According to the social force model, pedestrians demonstrate increased sensitivity to strangers within their field of view, resulting in a repulsion force. Conversely, they generate an additional attraction force for companions within their field of view. Notably, interest points and pedestrian trajectories outside their field of view do not exert influence on them. The POI-GAN model introduces a filtering approach based on the pedestrian field of view angle mechanism. This approach determines pedestrians' head angles, models the field of view range established by the predicted pedestrian, identifies the attention and focus ranges, and determines the impact of dynamic and static elements within the pedestrian's field of view on their trajectory.

Assuming pedestrians only focus on what lies ahead, at the initial moment t, the head rotation angle of the predicted pedestrian j can be represented as θ_t^j, According to research conducted by Luo et al. [35] and Atchison [36], the horizontal field of view attention range for this pedestrian is defined as approximately $\theta_t^{j,1} \approx 120°$, and the focus range is approximately $\theta_t^{j,2} \approx 30°$. The field of view domain of pedestrian j is expressed as follows:

$$w_t^j = \left\{(\rho, \sigma)|\rho \in (0, l], \sigma_t^j \in \left[\theta_t^j - \frac{\eta}{2}, \theta_t^j + \frac{\eta}{2}\right]\right\} \quad (8)$$

The operational states of the pedestrian field of view angle model can be classified as follows:

(1) In Figure 4(a), at time t, pedestrian m lies beyond the attention range of pedestrian n, while pedestrian n falls within the field of view of pedestrian m. Therefore, pedestrian n remains uninfluenced by pedestrian m when selecting a trajectory. Conversely, pedestrian m 's trajectory may be influenced by pedestrian n.

(2) In Figure 4(b), interest point i lies within the attention range of pedestrian m, and the Euclidean distance $r_t^{i,m}$ between them falls below r_2 but exceeds r_1. Consequently, pedestrian m is attracted toward interest point i.
(3) In Figure 4(c), interest point i is also falls within the attention range of pedestrian n but lies outside the focus range. Therefore, despite the Euclidean distance between them meeting the criteria $r_t^{i,n} < r_2$ and $r_t^{i,n} > r_1$, the trajectory of pedestrian n remains uninfluenced by interest point i.

E. IMPORTANCE SAMPLING ALGORITHM

Traditional pedestrian trajectory prediction research typically treats the attention of all pedestrians within the data collection range as equal. However, our study focuses more on pedestrians who may interact with points of interest. In certain scenarios, the number of pedestrians interacting with interest points is relatively small, resulting in low interaction frequencies. Consequently, the value of these interaction data surpasses the average value of all pedestrian trajectory data. If traditional uniform sampling and batch updating training methods are employed, it may lead to slow convergence of the model and poor prediction performance.

To address this challenge, the POI-GAN model utilizes the Annealing The Bias algorithm [37] during the training phase to enhance training efficiency. This method first segments the raw data evenly based on time intervals and assigns priority to each segment according to the frequency of interactions between pedestrians and points of interest within the segment. During training, the probability of extracting segments can be expressed as:

$$P(i) = \frac{p_i^\alpha}{\Sigma_k p_k^\alpha} \tag{9}$$

$$p_i = \frac{1}{\text{rank}(i)} \tag{10}$$

where:
α governs the extent of priority utilization, $\alpha = 0$ signifies uniform sampling, $\Sigma_k p_k^\alpha$ serves as a normalization factor to ensure that the sum of all $P(i)$ probabilities equals 1, rank(i) denotes the priority ranking of segment i.

To mitigate deviations, importance sampling and an annealing factor β are introduced. The sampling weight is expressed as:

$$w_i = \left(\frac{1}{N} \cdot \frac{1}{P(i)}\right)^\beta \tag{11}$$

where:
$\beta \in [0, 1]$, when $\beta = 0$, the importance sampling algorithm is not applied, while $\beta = 1$ signifies standard importance sampling.

At the beginning of training, the POI-GAN model sets β to 1, prioritizing the extraction of the extraction of high-priority samples. As training progresses, β gradually approaches 0, causing the importance sampling algorithm to progressively diminish until training is completed.

IV. EXPERIMENT AND ANALYSIS

A. EXPERIMENT SCENARIOS

In order to thoroughly investigate the performance of POI-GAN in terms of trajectory prediction and quantify its effectiveness in improving service quality in service scenarios, we designed two control experimental scenarios.

Experiment 1: We selected an outdoor stall scenario as the data collection site. This scenario comprises multiple stalls, which serve the dual purpose of attracting pedestrians to avail of services and acting as general obstacles that can hinder pedestrian passage, thus being categorized as points of interest. The service targets are the general public, and pedestrians who interact with the points of interest typically do not have a predefined intention before entering the scene, reflecting impulsive consumption behavior. The locations of the stalls and the types of goods sold are not fixed. In essence, pedestrians only actively approach and linger at specific stalls after expressing interest in them within their field of view. This experiment aims to illustrate that POI-GAN can accurately discern the attributes of points of interest as special obstacles that draw some pedestrians in public service scenarios, thereby offering trajectory predictions that are more rational compared to baseline models.

Experiment 2: We chose an indoor tearoom scenario as the data collection site. This scenario encompasses unique obstacles such as coffee machines, water dispensers, refrigerators, and snack trays, which are also regarded as points of interest. The service targets are fixed users, and the locations and service offerings of the points of interest remain constant. Pedestrians enter the scene with a clear intention to receive service; for instance, they may carry a cup to the water dispenser to obtain water, and they are unlikely to deviate from their walking route due to other points of interest within their field of view. This experiment aims to demonstrate that POI-GAN can forecast the automated services that pedestrians will receive in fixed-user service scenarios, thereby enabling service providers to prepare services in advance.

B. EXPERIMENTAL ENVIRONMENT AND TRAINING PARAMETER SETTINGS

Our experiments were conducted using consistent hardware conditions. The experimental setup included an AMD Ryzen R5-5600 CPU @3.5 GHz, featuring 6 cores and 12 threads, an NVIDIA 1080ti GPU with 11 GB of video memory, and utilized the PyTorch framework.

The generator was structured in an encoder-decoder, with the encoder comprising 128 LSTM units. The pooling module utilized 2 consecutive latent codes, a noise vector of length 62, and a pooling vector of size 64. Similarly, the decoder also consisted of 128 LSTM units, with weight sharing between LSTM layers performing identical functions. The discriminator utilized two LSTM blocks (LSTM-OE and LSTM-PE) with a hidden layer size of 128 for processing observed trajectories.

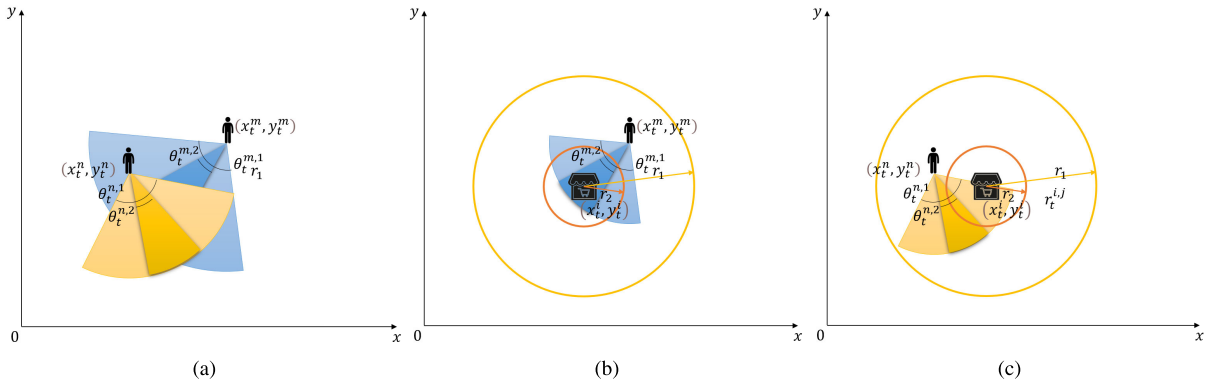


FIGURE 4. Interaction between points of interest and pedestrians.

Each dataset was trained using consistent hyperparameters for GAN network training, including a batch size of 64, a generator learning rate set to 0.001, and a discriminator learning rate set to 0.0001. The GAN training consisted of 20,000 rounds.

C. PUBLIC ORIENTED SERVICE SCENARIO EXPERIMENT

1) DATASETS AND EVALUATION METRICS

Pedestrian interaction video data was collected from a drone perspective, covering a rectangular area measuring 12 meters in length and 7.5 meters in width, as depicted in Figure 5. This region included 10 predefined interest points, labeled 1 through 10 in Figure 5, with lower numbers denoting interest points frequented by more visitors. Figure 6 exhibits the trajectories of select pedestrians over the previous 3.2 seconds, while Figure 7 depicts the center line of the pedestrian field of view.

FIGURE 5. Data collection scenario.

In the experiment, we selected 17 segments and annotated 1840 key frames at intervals of 0.4 seconds. Pedestrians' two-dimensional coordinates were recorded frame by frame, adhering to the format of the ETH and UCY datasets. Additionally, we annotated the deflection angle θ of the pedestrian field of view centerline. We allocated 80% of the data as the training and validation set, with the remaining

FIGURE 6. Partial pedestrian trajectories.

20% designated as the test set. The data format is presented in Table 1.

TABLE 1. Data format.

Frame	Pedestrian	Position X	Position Y	Deflection Angle
300	12	12.55	5.30	130
301	12	12.70	5.35	135

Consistent with prior research methodologies [14], [15], [38], our evaluation metrics included ADE (average displacement error) and FDE (final displacement error). ADE quantified the accuracy of predictive sequences by determining the mean Euclidean distance between the predicted trajectory and the actual trajectory at each time step. Meanwhile, FDE assessed predictive sequence accuracy by measuring the mean Euclidean distance between the predicted trajectory position and the actual trajectory position at the final time step. Additionally, we explored the correlation between training iterations and trajectory errors, along with the

FIGURE 7. Pedestrian deflection angle.

influence of prediction duration on service hit rate, to gauge the algorithm's practical viability.

The comparative analysis included the following models, each briefly summarized as follows:

- LSTM [14]: A basic LSTM model without interactions between pedestrians or with the environment.
- Social LSTM [14]: Modeled each pedestrian with a separate LSTM, aggregating hidden states using a social pooling layer at each time step.
- Social GAN [38]: Utilized a generative adversarial network applied to the S-LSTM prediction model.
- Social Ways [15]: Employed a generative pedestrian trajectory prediction model based on Info-GAN and an attention mechanism.
- Social-BiGAT [39]: A discriminative model that introduces bidirectional graph attention mechanisms.
- Social-STGCNN [19]: A discriminative model based on graph convolutional neural networks (GCNN).

2) PERFORMANCE COMPARISON

Table 2 presents a comparative assessment of ADE and FDE metrics between POI-GAN and six different trajectory prediction algorithms at prediction durations of 8 frames (3.2 seconds) and 12 frames (4.8 seconds). The scene comprises 5 interest points, delineated as 1, 2, 3, 4, and 5 in Figure 4.

In terms of the FDE_8 and FDE_{12} metrics, POI-GAN demonstrates significant improvements relative to other methods. This enhancement primarily stems from the introduction of the point of interest (POI) model into the scene, allowing the model to accurately understand the interactions between pedestrians and points of interest, rather than simply treating points of interest as generic obstacles. Consequently, POI-GAN exhibits superior performance in predicting the endpoints of trajectories for these pedestrians.

In contrast, the LSTM algorithm treats each pedestrian as an independent entity, lacking interaction information among pedestrians, thus resulting in relatively larger errors. The Social LSTM algorithm models pedestrian interactions on top of LSTM, rendering the generated pedestrian trajectories more reasonable and significantly reducing errors. Both Social GAN and Social Ways are generative models based on GAN and LSTM, with broader social pooling scopes, thereby enabling better understanding of global information and further enhancing prediction performance. On the other hand, Social-BiGAT and Social-STGCNN adopt graph-based approaches, enabling better capture of spatial and temporal relationships between trajectories. Particularly, owing to its adaptability to complex social scenes, Social-STGCNN performs exceptionally well in crowded environments, thereby highlighting the prominence of the ADE_8 and ADE_{12} metrics.

TABLE 2. Comparison of trajectory prediction algorithm errors.

Metrics	Trajectory Prediction Algorithm						
	LSTM	Social LSTM	Social GAN	Social Ways	Social-BiGAT	Social-STGCNN	POI-GAN
ADE_8	0.98	0.70	0.54	0.42	0.42	**0.39**	0.39
FDE_8	1.81	1.13	0.72	0.77	0.73	0.69	**0.56**
ADE_{12}	1.33	1.12	1.03	0.80	0.74	**0.67**	0.71
FDE_{12}	2.10	1.65	1.49	1.36	1.27	1.13	**1.07**

3) SENSITIVITY ANALYSIS

Sensitivity analyses were carried out to verify the efficacy of the Point of Interest (POI) model and the field of view (FOV) model within the POI-GAN framework. We selected Social Ways as the baseline method and conducted comparative experiments by incorporating the FOV model, interest point model, and a hybrid of both into the baseline. The consolidated findings are outlined in Table 3.

The experiments illustrate that the incorporation of the FOV model results in marginal enhancements across several metrics, highlighting its ability to effectively screen out illogical interaction behaviors. Furthermore, the integration of the interest point model improves the predictiveness of endpoints in certain pedestrian trajectories, notably reflected in the notable enhancements observed in the FDE_8 and FDE_{12} metrics.

TABLE 3. POI-GAN sensitivity analysis.

Metrics	Trajectory Prediction Algorithm			
	Base	Base+FOV	Base+POI	Base+FOV+POI
ADE_8	0.42	0.39	0.40	0.38
FDE_8	0.77	0.71	0.62	0.56
ADE_{12}	0.80	0.74	0.74	0.71
FDE_{12}	1.36	1.32	1.17	1.07

To evaluate the efficacy of the importance sampling algorithm employed during the training phase of POI-GAN, we investigated the correlation between training iterations

and the ADE and FDE metrics with and without the utilization of the importance sampling algorithm. The experimental findings, depicted in Figure 8, indicate that with the implementation of the importance sampling algorithm (Annealing The Bias), the model achieved comparable outcomes to conventional training with 300 iterations, as opposed to the standard 800 iterations. As the number of training iterations increased, the importance sampling algorithm consistently outperformed in terms of ADE and FDE metrics. This underscores the algorithm's ability to improve training efficiency. Remarkably, the algorithm enables substantial reductions in hardware resource consumption while maintaining training efficacy, rendering it particularly beneficial for hardware-limited platforms.

4) CASE ANALYSIS

In Figure 9(a), we depict trajectory prediction outcomes for pedestrians $p1$ and $p2$ in the same scene and timeframe, generated by the POI-GAN model and the Social Ways model. In the illustration, solid lines and circles represent the actual trajectories of the pedestrians, squares depict the predicted trajectories of the POI-GAN model, and triangles illustrate the predicted trajectories of the Social Ways model. A comparison of these results reveals an intriguing distinction: the Social Ways model presupposes mutual influence between pedestrians in the front and rear positions, with the front pedestrian making evasive maneuvers to avoid the one behind. Conversely, the POI-GAN prediction indicates that the trajectory of the front pedestrian remains unaffected by the pedestrian behind, while the latter's trajectory is indeed influenced by the former, demonstrating closer alignment with real-world scenarios.

Figure 9(b) presents another set of comparative samples under the same experimental conditions. The Social Ways model regards the interest points within the scene as impassable and uninviting obstacles, predicting that pedestrians would choose to detour around them. In contrast, the POI-GAN model adeptly foresees pedestrians pausing and lingering near these interest points.

5) PARAMETER ANALYSIS

To evaluate the impact of the quantity of predefined interest points within the scene on trajectory prediction accuracy, we performed a comparative analysis of ADE and FDE indicators across various quantities of interest points quantities ranging from 0 to 10. Figure 5 illustrates the coordinates of these interest points, prioritizing those with lower numerical designations when the count is below 10.

The experimental findings depicted in Figure 10, demonstrate a significant influence of the interest point quantity parameter on the FDE_8 and FDE_{12} indicators, with optimal outcomes observed when the number of interest points is set to 5. Through analysis and discussion of the trajectories generated by the model, we believe that when the number of predefined interest points is less than 5, the representation ability of POI-GAN for pedestrian trajectories near interest points is insufficient, resulting in a higher number of misjudged trajectories around undefined interest points. Conversely, when the number of interest points exceeds 5, the interaction frequency between larger serial number interest points and surrounding pedestrians is lower, leading to insufficient training data for the model and a decrease in prediction accuracy. Notably, interest points labeled as 7, 8, and 10 exhibit strip-like configurations, where the social force interest point model employed by POI-GAN shows suboptimal performance.

D. EXPERIMENT IN SERVICE SCENARIOS TARGETING FIXED AUDIENCE

1) DATASET CREATION

To address the issues of high equipment costs and expensive manual labeling in the process of creating trajectory datasets based on image collection, we propose an innovative data collection scheme. This scheme utilizes the HLK-LD6001A-60G millimeter-wave radar as the data source, integrates it with the HomeAssistant platform through an ESP32 gateway, and leverages the Node-Red plugin to interact with third-party smart devices (such as Xiaomi Mi Home). This end-to-end automation service scheme enables POI-GAN to efficiently utilize data for training and testing. Figure 11 illustrates the overall framework of this scheme.

In this scheme, pedestrian trajectory data is obtained from the HLK-LD6001A-60G millimeter-wave radar, which can track the positions of human bodies within a square area with a side length of 8m at a rate of 30 frames per second and can simultaneously track 8 targets, thus providing direct access to two-dimensional coordinates of pedestrians within the area. The radar outputs data to the ESP32 via serial port. When the ESP32 is on the same local area network as the HomeAssistant server, data can be accessed through the IP address and port number into the Node-Red plugin installed on the HomeAssistant platform. During data collection, Node-Red extracts pedestrian IDs and coordinates based on a pre-defined decoding protocol, adds timestamps, and saves them to log files. Additionally, Node-Red supports the integration of other third-party devices. By capturing data sent by these devices, we can clearly understand whether users have accepted a service and when they have received the service. This data can also be saved to log files and analyzed in conjunction with radar data.

The innovation of this scheme lies in the use of millimeter-wave radar as a low-cost data collection device, avoiding the cost and technical barriers associated with traditional image collection methods. Moreover, through integration with smart home platforms, we can obtain richer and more diverse data, providing more possibilities for the training and application of trajectory prediction models.

In the experiment targeting fixed audience service scenarios, we adopted a data collection scheme based on millimeter-wave radar, with a tearoom as the collection scene. We selected a rectangular area measuring 6m in length and 4.5m in width as the collection area, as shown in Figure 12. When the door leading to the office area and balcony is open, the radar signal can extend through the

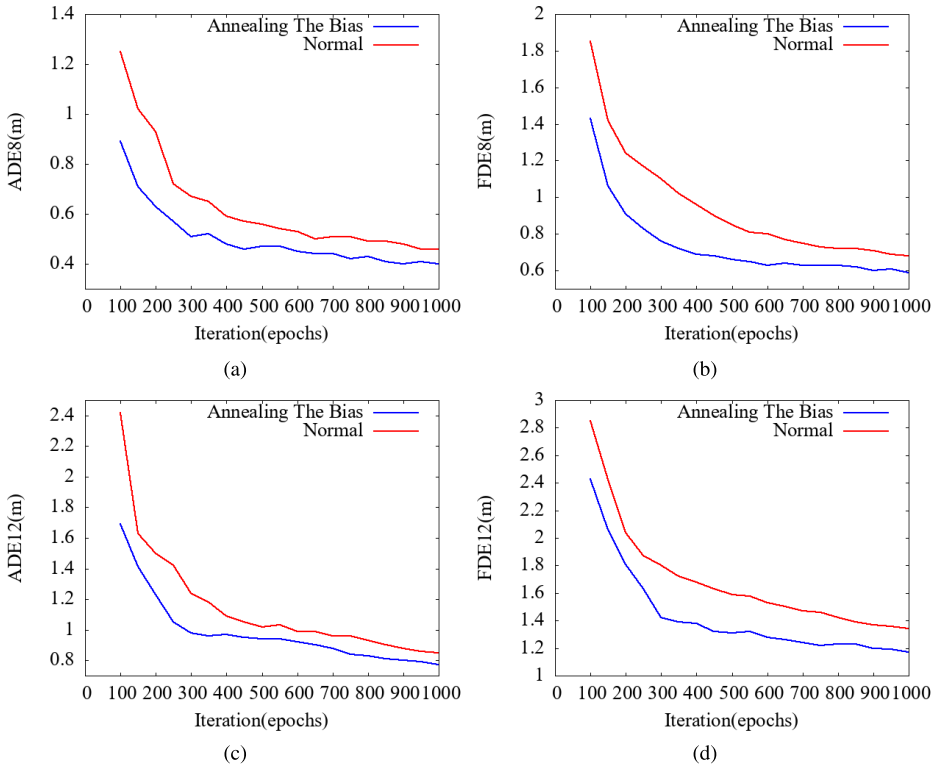


FIGURE 8. Comparative analysis of iteration times and training errors of importance sampling algorithm.

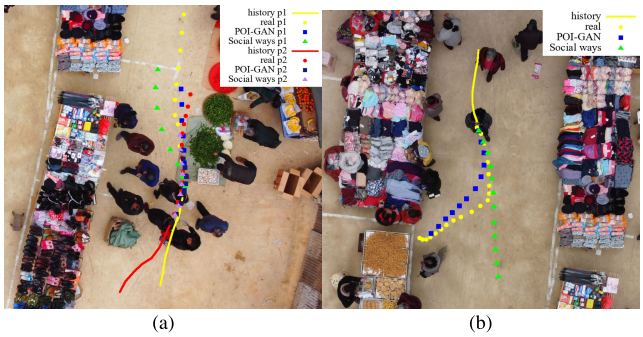


FIGURE 9. Comparison of POI-GAN and social ways generation trajectories.

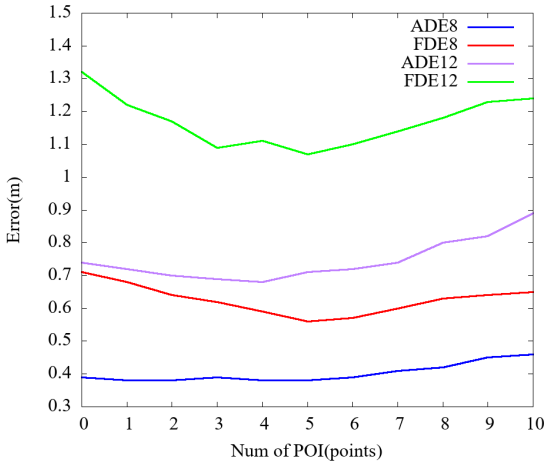


FIGURE 10. Analysis of the relationship between the number of POI-GAN interest points and prediction error.

door, enabling the POI-GAN model to obtain pedestrian coordinates earlier and make trajectory predictions, thereby providing more preparation time for automated services. The collection area includes four predefined points of interest, labeled as numbers 1 through 4.

Among them, point of interest 1 is a coffee machine, which serves as an automated service provider capable of providing self-service hot drinks for office staff. Each time a user uses the coffee machine, they need to press the preheat button on the top of the machine to preheat the water line. The preheating time depends on the interval between machine uses, typically ranging from 3 to 10 seconds during weekday office hours. Our experimental goal is to use pedestrian trajectory prediction to prompt the coffee machine to preheat in advance, thereby reducing the waiting time for users standing in front of the coffee machine. The power socket of the coffee machine is connected to a Xiaomi smart power plug, which can be integrated into the HomeAssistant platform, reporting real-time switch status and power consumption data. By monitoring power changes, we can infer the usage status of the coffee machine, such as standby, preheating, and dispensing.

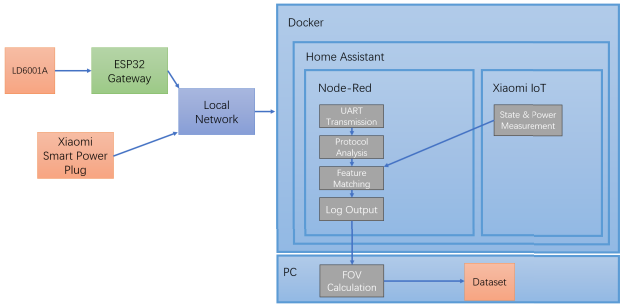


FIGURE 11. Analysis of the relationship between POI-GAN service hit rate and prediction time.

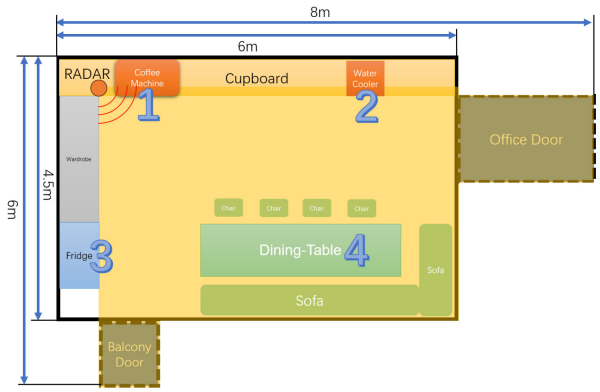


FIGURE 12. Indoor data collection scenario.

We collected raw data for 15 working days, totaling 120 hours, using millimeter-wave radar and the HomeAssistant platform. After excluding invalid data records when the indoor population exceeded 8 people during lunchtime, we obtained approximately 1.8K valid pedestrian trajectory data points. Since the millimeter-wave radar cannot perceive the turning angle of pedestrians' heads, we used the direction of pedestrian movement instead of head turning angle in this experiment. When the pedestrian's movement speed is less than 0.5m/s, the field of view module does not operate, and the deviation angle of the central line in the field of view is recorded as -1. The data format is the same as that described in Section IV-C, as shown in Table 1. We selected 80% of the data as the training and validation sets and 20% of the data as the test set.

To assess the benefits brought by various optimization schemes to automated services, we define a profit function. The total profit brought by the model for automated services can be represented as:

$$V_P = V_B(H + E + M) + V_H \cdot H - V_E \cdot E - V_F \quad (12)$$

where V_B represents the basic profit provided by the automated service itself, V_H represents the additional profit brought by the optimization algorithm when it hits, V_E represents the loss brought by the optimization algorithm when it misjudges, and V_F represents other overhead costs introduced by the optimization algorithm, with single time overhead not included. Here, H represents the number of hits, E represents the number of misjudgments, and M is regarded as a missed hit with no additional profit or loss.

In this experiment, we will use the following optimization schemes for profit comparison. Here is a brief introduction to each scheme:

Base: No optimization is used, and each cup of coffee provided counts as one basic profit $V_B = 1$.

Part-Time: During working hours on each workday, the coffee machine is kept in a preheated state for an extended period, considered as 100% hits. For each cup of coffee provided, an additional hit profit $V_H = 0.4$ is added on top of the basic profit. Additional power consumption and machine wear due to prolonged high-temperature operation introduce additional cost $V_F = 10$.

POI-GAN-2.4S: POI-GAN is used for trajectory prediction with a prediction duration of 2.4 seconds. When the predicted trajectory of a pedestrian within the next 2.4 seconds overlaps with the range of the coffee machine point of interest and the trajectory endpoint is also within this range, the coffee machine is commanded to preheat in advance. For each hit, an additional hit profit $V_H = 0.18$ is added on top of the basic profit, and a misjudgment introduces a loss $V_E = 0.02$.

POI-GAN-3.2S: Similar to POI-GAN-2.4S, but with a prediction duration of 3.2 seconds, hit profit $V_H = 0.24$, and misjudgment loss $V_E = 0.02$.

POI-GAN-3.2S-N: Based on POI-GAN-3.2S, the field of view module is removed to evaluate the feasibility of using the direction of movement instead of the central line of the field of view in such scenarios, with no change in hit profit and misjudgment loss.

POI-GAN-3.2S-O: POI-GAN is used for trajectory prediction with a prediction duration of 3.2 seconds. Each prediction generates 20 trajectories, and if one or more of these trajectories overlap with the range of the coffee machine point of interest and the trajectory endpoint is also within this range, the coffee machine is commanded to preheat in advance. Hit profit and misjudgment loss are the same as POI-GAN-3.2S.

POI-GAN-4.0S: Similar to POI-GAN-2.4S, but with a prediction duration of 4.0 seconds, hit profit $V_H = 0.27$, and misjudgment loss $V_E = 0.02$.

POI-GAN-4.8S: Similar to POI-GAN-2.4S, but with a prediction duration of 4.8 seconds, hit profit $V_H = 0.3$, and misjudgment loss $V_E = 0.02$. Considering the limited perception range of the millimeter-wave radar, the actual effective prediction duration for pedestrians to approach the coffee machine may be less than 4.8 seconds.

2) PERFORMANCE COMPARISON

We compare the Base scheme with the other six optimization schemes in terms of profit and cost, and summarize the results in Table 4:

Through a detailed analysis of the experimental results, we draw the following conclusions:

(1) When the coffee machine is used frequently, maintaining preheating for an extended period significantly increases profits compared to the baseline scheme. However,

TABLE 4. Profit and cost comparison.

Scheme	Evaluation Metrics				
	Base Profit	Hit Profit	Miss Loss	Additional Cost	Total Profit
Base	644	0	0	0	644
Part-Time	644	257.60	0	150	751.60
POI-GAN-2.4S	644	97.38	1.68	0	739.70
POI-GAN-3.2S	644	122.16	3.06	0	763.10
POI-GAN-3.2S-N	644	118.80	4.62	0	758.18
POI-GAN-3.2S-O	644	151.44	11.73	0	783.71
POI-GAN-4.0S	644	125.28	4.98	0	764.30
POI-GAN-4.8S	644	134.40	5.94	0	772.46

this requires optimization of the machine design to ensure long-term operation without negative impacts, such as improving insulation efficiency and ensuring water quality safety.

(2) In practical applications, choosing a longer prediction time helps increase hit profits. However, when the prediction time exceeds the time required for the coffee machine to preheat, the surplus prediction time only wastes computing power without bringing additional profits. Therefore, the prediction duration needs to be carefully considered based on the actual situation.

(3) POI-GAN is suitable for automated services with significant hit profits and minimal misjudgment losses, such as the coffee machine preheating service chosen in this experiment, where each misjudgment only incurs a small amount of power consumption. Furthermore, as a generative model, POI-GAN can output multiple predicted trajectories for a single target, further increasing the model's recall rate and improving service quality.

(4) For office tea rooms, where the recipients of automated services are fixed users whose behavior is less affected by other points of interest and pedestrians within the field of view, the effect of the field of view module is not significant in this experiment. These characteristics of the scene minimize the drawbacks of the millimeter-wave radar solution's inability to detect the turning angles of pedestrians' heads.

(5) Due to the spatial structure of the tearoom and the deployment position of the millimeter-wave radar, the hit rate of POI-GAN significantly decreases when the prediction duration exceeds 4.0 seconds. Therefore, the application of POI-GAN is limited in scenarios with narrow spaces and dense points of interest.

V. CONCLUSION

POI-GAN is a pedestrian trajectory prediction method that leverages the Point of Interest (POI) model and the field of view angle model. By employing generative adversarial networks, it produces diverse pedestrian prediction trajectories, showcasing enhanced prediction accuracy particularly in the presence of points of interest within the scene. The incorporation of the POI model enables POI-GAN to accurately identify special obstacles in the scene, thereby effectively improving the accuracy of predicting trajectory endpoints. Furthermore, the field of view angle model serves to filter out unrealistic interaction behaviors, consequently yielding more realistic generated trajectories. Moreover, the utilization of the biased annealing algorithm during model training aids in the learning of high-value trajectories from a limited subset of samples, thereby enhancing training efficiency.

However, due to limitations in the social force model used for modeling points of interest, POI-GAN is not suitable for scenes with very high crowd densities where queueing phenomena may occur [40]. Moreover, constrained by data dimensions, POI-GAN may struggle to identify specific interaction targets when multiple points of interest are closely spaced. Furthermore, POI-GAN is not adept at handling circular, linear, or movable points of interest that may exist in real-world scenarios. Future work will focus on optimizing point of interest modeling to enhance the model's applicability.

REFERENCES

[1] G. Medberg and C. Grönroos, "Value-in-use and service quality: Do customers see a difference?" *J. Service Theory Pract.*, vol. 30, no. 4/5, pp. 507–529, Nov. 2020.

[2] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, "RFM ranking—An effective approach to customer segmentation," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 33, no. 10, pp. 1251–1257, Dec. 2021.

[3] N. Sharma, C. Dhiman, and S. Indu, "Pedestrian intention prediction for autonomous vehicles: A comprehensive survey," *Neurocomputing*, vol. 508, pp. 120–152, Oct. 2022.

[4] D. Ridel, E. Rehder, M. Lauer, C. Stiller, and D. Wolf, "A literature review on the prediction of pedestrian behavior in urban scenarios," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 3105–3112.

[5] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, Jul. 2020.

[6] P. Kothari, S. Kreiss, and A. Alahi, "Human trajectory forecasting in crowds: A deep learning perspective," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7386–7400, Jul. 2022.

[7] F. Camara, N. Bellotto, S. Cosar, F. Weber, D. Nathanael, M. Althoff, J. Wu, J. Ruenz, A. Dietrich, G. Markkula, A. Schieben, F. Tango, N. Merat, and C. Fox, "Pedestrian models for autonomous driving—Part II: High-level models of human behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5453–5472, Sep. 2021.

[8] B. I. Sighencea, R. I. Stanciu, and C. D. Căleanu, "A review of deep learning-based methods for pedestrian trajectory prediction," *Sensors*, vol. 21, no. 22, p. 7543, Nov. 2021.

[9] S. Kolekar, S. Gite, B. Pradhan, and K. Kotecha, "Behavior prediction of traffic actors for intelligent vehicle using artificial intelligence techniques: A review," *IEEE Access*, vol. 9, pp. 135034–135058, 2021.

[10] R. Korbmacher and A. Tordeux, "Review of pedestrian trajectory prediction methods: Comparing deep learning and knowledge-based approaches," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 24126–24144, Dec. 2022.

[11] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 51, no. 5, pp. 4282–4286, May 1995.

[12] A. Alahi, V. Ramanathan, and L. Fei-Fei, "Socially-aware large-scale crowd forecasting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2211–2218.

[13] S. Yi, H. Li, and X. Wang, "Understanding pedestrian behaviors from stationary crowd groups," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3488–3496.

[14] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 961–971.

[15] J. Amirian, J.-B. Hayet, and J. Pettré, "Social ways: Learning multi-modal distributions of pedestrian trajectories with GANs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 2964–2972.

[16] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.

[17] S. Liu, H. Liu, H. Bi, and T. Mao, "CoL-GAN: Plausible and collision-less trajectory prediction by attention-based GAN," *IEEE Access*, vol. 8, pp. 101662–101671, 2020.

[18] L. Huang, J. Zhuang, X. Cheng, R. Xu, and H. Ma, "STI-GAN: Multimodal pedestrian trajectory prediction using spatiotemporal interactions and a generative adversarial network," *IEEE Access*, vol. 9, pp. 50846–50856, 2021.

[19] A. Mohamed, K. Qian, M. Elhoseiny, and C. Claudel, "Social-STGCNN: A social spatio-temporal graph convolutional neural network for human trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 14412–14420.

[20] C. Choi, S. Malla, A. Patil, and J. H. Choi, "DROGON: A trajectory prediction model based on intention-conditioned behavior reasoning," 2019, *arXiv:1908.00024*.

[21] Y. Xu, L. Wang, Y. Wang, and Y. Fu, "Adaptive trajectory prediction via transferable GNN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 6510–6521.

[22] X. Mo, Z. Huang, Y. Xing, and C. Lv, "Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9554–9567, Jul. 2022.

[23] B. Varadarajan, A. Hefny, A. Srivastava, K. S. Refaat, N. Nayakanti, A. Cornman, K. Chen, B. Douillard, C. P. Lam, D. Anguelov, and B. Sapp, "MultiPath++: Efficient information fusion and trajectory aggregation for behavior prediction," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 7814–7821.

[24] R. W. Liu, M. Liang, J. Nie, Y. Yuan, Z. Xiong, H. Yu, and N. Guizani, "STMGCN: Mobile edge computing-empowered vessel trajectory prediction using spatio-temporal multigraph convolutional network," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 7977–7987, Nov. 2022.

[25] X. Jia, P. Wu, L. Chen, Y. Liu, H. Li, and J. Yan, "HDGT: Heterogeneous driving graph transformer for multi-agent trajectory prediction via scene encoding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 11, pp. 13860–13875, Nov. 2023.

[26] T. Gilles, S. Sabatini, D. Tsishkou, B. Stanciulescu, and F. Moutarde, "GOHOME: Graph-oriented heatmap output for future motion estimation," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 9107–9114.

[27] C. Xu, M. Li, Z. Ni, Y. Zhang, and S. Chen, "GroupNet: Multiscale hypergraph neural networks for trajectory prediction with relational reasoning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 6488–6497.

[28] A. Kalatian and B. Farooq, "A context-aware pedestrian trajectory prediction framework for automated vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 134, Jan. 2022, Art. no. 103453.

[29] Z. Lv, X. Huang, and W. Cao, "An improved GAN with transformers for pedestrian trajectory prediction models," *Int. J. Intell. Syst.*, vol. 37, no. 8, pp. 4417–4436, Aug. 2022.

[30] C. Xu, W. Mao, W. Zhang, and S. Chen, "Remember intentions: Retrospective-memory-based trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 6478–6487.

[31] S. Liu, S. Tripathi, S. Majumdar, and X. Wang, "Joint hand motion and interaction hotspots prediction from egocentric videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 3272–3282.

[32] Z. Chen, Q. Guo, T. Li, Y. Yan, and D. Jiang, "Gait prediction and variable admittance control for lower limb exoskeleton with measurement delay and extended-state-observer," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 18, 2022, doi: 10.1109/TNNLS.2022.3152255.

[33] N. Stein, G. Bremer, and M. Lappe, "Eye tracking-based LSTM for locomotion prediction in VR," in *Proc. IEEE Conf. Virtual Reality 3D User Interface (VR)*, Mar. 2022, pp. 493–503.

[34] B. Xia, C. Wong, Q. Peng, W. Yuan, and X. You, "CSCNet: Contextual semantic consistency network for trajectory prediction in crowded spaces," *Pattern Recognit.*, vol. 126, Jun. 2022, Art. no. 108552.

[35] X. Luo, X. He, L. Qing, X. Chen, L. Liu, and Y. Xu, "EyesGAN: Synthesize human face from human eyes," *Neurocomputing*, vol. 404, pp. 213–226, Sep. 2020.

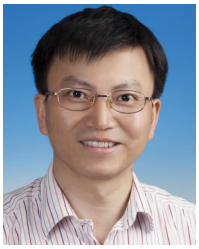
[36] D. Atchison, *Optics of the Human Eye*. Boca Raton, FL, USA: CRC Press, 2023.

[37] Y. Hou, L. Liu, Q. Wei, X. Xu, and C. Chen, "A novel DDPG method with prioritized experience replay," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 316–321.

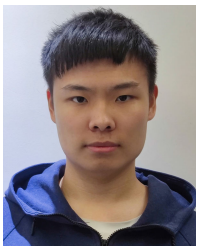
[38] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2255–2264.

[39] V. Kosaraju, A. Sadeghian, R. Martín-Martín, I. Reid, H. Rezatofighi, and S. Savarese, "Social-BiGAT: Multimodal trajectory forecasting using bicycle-GAN and graph attention networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–10.

[40] A. Vemula, K. Muelling, and J. Oh, "Social attention: Modeling attention in human crowds," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 4601–4607.



YE LI received the Ph.D. degree from the School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, in 2007. He is currently a Senior Engineer with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology. His research interests include machine learning, image processing, computer vision, and mobile communications.



CHI ZHANG was born in Yancheng, Jiangsu, China, in 1998. He received the B.S. degree from the City Institute, Dalian University of Technology, Dalian, China, in 2020. He is currently pursuing the M.S. degree with the University of Shanghai for Science and Technology. His research interests include pedestrian trajectory prediction, cache strategy, and edge computing.

JINGKANG ZHOU was born in Zhengzhou, Henan, China, in 2000. He received the B.S. degree from Shanghai Polytechnic University, Shanghai, China, in 2021. He is currently pursuing the M.S. degree with the University of Shanghai for Science and Technology. His research interests include deep learning image classification, semi-supervised learning, pattern recognition, and prediction.

SHENGCUI ZHOU was born in Jining, Shandong, China, in 1998. She received the B.S. degree from Shandong Jiaotong University, in 2020. She is currently pursuing the M.S. degree in electronic information with the University of Shanghai for Science and Technology, Shanghai, China. Her research interests include deep learning and image processing.