From Cognition to Precognition: A Future-Aware Framework for Social Navigation

Zeying Gong[1], Tianshuai Hu[2], Ronghe Qiu[1] and Junwei Liang[1,2*]

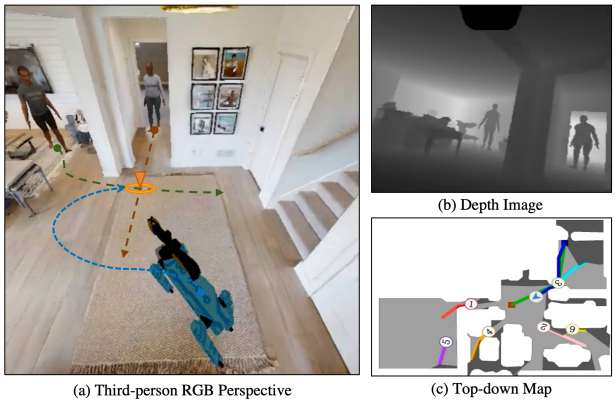
***Abstract*— To navigate safely and efficiently in crowded spaces, robots should not only perceive the current state of the environment but also anticipate future human movements. In this paper, we propose a reinforcement learning architecture, namely *Falcon*, to tackle socially-aware navigation by explicitly predicting human trajectories and penalizing actions that block future human paths. To facilitate realistic evaluation, we introduce a novel SocialNav benchmark containing two new datasets, Social-HM3D and Social-MP3D. This benchmark offers large-scale photo-realistic indoor scenes populated with a reasonable amount of human agents based on scene area size, incorporating natural human movements and trajectory patterns. We conduct a detailed experimental analysis with the state-of-the-art learning-based method and two classic rule-based path-planning algorithms on the new benchmark. The results demonstrate the importance of future prediction and our method achieves the best task success rate of 55% while maintaining about 90% personal space compliance. We will release our code and datasets.**

I. INTRODUCTION

Social navigation (SocialNav) refers to autonomous robot adhereing to *social norms* and *social etiquette* while navigating environments shared with humans [1]. This task poses new challenges to visual navigation, as modular approaches relying on pre-built maps struggle in dynamic, human-populated environments where collision avoidance is crucial.

Existing RL-based approaches [2], [3], [4] often struggle with dynamic environments due to limited foresight and reliance on global information [5], [6], [7]. Consider the example in Fig. 1, where a robot navigates to a goal located at the intersection of two nearby humans' future paths. In such cases, traditional RL approaches may struggle to avoid humans due to limited foresight or reliance on global information. In contrast, our approach addresses these issues by explicitly predicting human trajectories that enable social compliance and long-term dynamic collision avoidance.

Human trajectory forecasting has been shown to improve collision avoidance and navigation in dynamic environments [8], [9], [10], but it is mostly applied in outdoor scenarios like autonomous driving [11], [12], [13]. This motivates us to integrate trajectory prediction algorithms into the SocialNav task. Indoor environments pose additional challenges, such as limited space and maneuverability, increasing the risk of collisions [7]. To address these challenges, we develop ***Falcon***, a future-aware SocialNav framework with precognition. The key novelties of *Falcon* lie in its future awareness. First, it introduces the **Social Cognition Penalty**, including a trajectory obstruction penalty, encouraging the agent to proactively avoid potential collisions and follow social etiquette. Second, it employs a **Spatial-Temporal Precognition Module** that incorporates socially-aware auxiliary tasks, including trajectory prediction, to enhance the agent's comprehension of future dynamics during training.



(a) Third-person RGB Perspective (b) Depth Image (c) Top-down Map

Fig. 1: We integrate trajectory prediction into the SocialNav task. In (a), the robot navigates toward a goal while predicting human trajectories (dashed lines) and avoiding them, following social etiquette. The robot uses depth input as shown in (b). (c) offers a top-down map for reference, which is not used by the robot.

Another challenge in SocialNav is the lack of realism in configurations [14]. Current approaches typically simplify the environment to only include a robot and surrounding humans, neglecting the complexity of the scene itself [15], [16], [17]. Moreover, the solutions often assume the robot has access to global information, such as real-time human positions or a full map of the environment [18]. To overcome these unrealistic settings, we introduce a novel SocialNav benchmark that contains two new datasets, **Social-HM3D** and **Social-MP3D**. We construct the datasets using 3D-reconstructed real-world indoor scenes. The scenes are populated with collision-avoidant humans moving toward their goals. The robot is given only one point goal and egocentric inputs during inference, without relying on a global map or known human trajectories. Our benchmark offers a more realistic representation of social navigation. In addition, we verify *Falcon* on the proposed benchmark, achieving state-of-the-art performance with a 55% success rate while maintaining a high social compliance score.

Overall, our main contributions are as follows:

- We introduce the first realistic SocialNav benchmark with two novel datasets, **Social-HM3D** and **Social-**

[1] The Hong Kong University of Science and Technology (Guangzhou). {zgong313,rqiu683}@connect.hkust-gz.edu.cn, junweiliang@hkust-gz.edu.cn * Corresponding author.
[2] The Hong Kong University of Science and Technology. thuaj@connect.ust.hk

MP3D, featuring large-scale photo-realistic scenes with realistic human and robot animations.

- We propose a new and effective SocialNav framework, ***Falcon***, that integrates explicit trajectory prediction, allowing the robot to perceive and predict for safe, socially comfortable and effective navigation.
- We establish a new state-of-the-art result compared to prior approaches on the proposed benchmark.

II. Related Works

A. Social Navigation.

In this paper, we focus on the SocialNav task [7], [14], which was first introduced in the iGibson SocialNav Challenge [19], building on the PointGoal Navigation (PointNav) by adding moving humans. Humans in the challenge are static figures with unrealistic movements. In contrast, our work utilizes the recent Habitat 3.0 simulator [20] to leverage realistic human movements and animations, despite our task differs from the SocialNav task as defined in Habitat 3.0.

SocialNav has been widely studied in robotics, computer vision, and social behavior analysis [21], [22]. In collision-free multi-agent navigation [23], [24], [25] and dynamic environments [26], research has advanced to address challenges posed by the presence of humans [27], [28], [29], [30], [31]. Other approaches [28], [31] model human-agent interactions through spatio-temporal graphs [30] to capture agent dynamics over time. Recent studies have also focused on egocentric SocialNav in photo-realistic or real-world scenarios [32], [33], [34]. Differing from these approaches, our method introduces explicit trajectory prediction within auxiliary tasks to train an RL-based agent for SocialNav.

B. Auxiliary tasks in Navigation.

Auxiliary tasks have proven effective for improving navigation performance through self-supervised learning [35]. While existing methods focus on environmental property prediction or agent state estimation [36], [37], [38], our framework introduces socially-aware auxiliary tasks that go beyond static observations. By predicting human future trajectories, we enable agents to proactively anticipate and comply with social norms, a critical requirement for dynamic indoor environments [18].

C. Human Trajectory Prediction.

Human trajectory prediction is vital for enabling safe and intelligent behavior in autonomous systems [39], [40]. Traditional approaches often rely on physical models, such as the Social Force model [28], which uses attractive and repulsive forces to simulate social behaviors and collision avoidance. Broadly, these methods fall into three categories. One approach is based on physics, where explicit dynamical models are derived from Newton's laws of motion to predict trajectories [41], [42], [43]. Another set of methods focuses on learning motion patterns from observed historical trajectories [44], [45], [46]. Moreover, planning-based methods aim to reason about the motion intent of rational agents, predicting trajectories by understanding agents' goals and their decision-making processes [47], [48], [49]. Our approach leverages these insights to develop a method that not only predicts human trajectories but also integrates socially-aware information into the agent's navigation policy, ensuring safe and efficient navigation in dynamic scenes.

III. Methodology

A. Problem Formulation

Consider a social navigation task where a robot a navigates in an environment populated by N active humans, $i \in \{1, \dots, N\}$. Starting from an initial configuration $q_a \in Q$, the robot aims to continuously select actions to generate a path τ_a toward its goal configuration $g_a \in Q$ while avoiding collision with static obstacles and dynamic humans. The overall objective is formulated as follows:

$$\begin{aligned} \tau_a = \arg\min_{\tau \in \mathcal{T}} & \left(c_a(\tau) + \lambda_a c_a^s(\tau, \tau_{1:N})\right) \\ \text{s.t.} \quad & A_a(\tau_a) \notin C_{\text{obs}}, \quad A_a(\tau_a) \cap A_i(\tau_i) = \emptyset, \\ & \tau_a(0) = q_a, \quad \tau_a(T) = g_a \end{aligned} \tag{1}$$

where c_a is the path cost guiding the robot to its goal; c_a^s is the cost term accounts for social norms; $A(\tau)$ represents the volume occupied along trajectory τ; C_{obs} represents static obstacles; T is the episode end time; and λ_a is a weight factor. The constraints make sure the robot does not collide with static obstacles and humans before reaching its goal.

Fig. 2 is an overview of *Falcon*. The **Main Policy Network** takes a depth image and a point goal at each timestep as inputs and directly outputs the robot's actions for the next step. It is trained using PointNav rewards along with our proposed **Social Cognition Penalty (SCP)**. In addition, the network is accompanied by the **Spatial-Temporal Precognition Module** that facilitates several auxiliary tasks during training. We detail each module in the subsequent sections.

B. Main Policy Network

The main policy network has two key components: the state encoders that extract visual and temporal features from observations, and a social cognition penalty that encourages social compliance. At each timestep, the point goal is transformed by a linear encoder whereas the depth image is processed by a ResNet-50 [50] visual encoder. A 2-layer LSTM [51] extracts the temporal features, which are then input to the actor head to produce actions and a value head to predict rewards. The LSTM hidden state is also used as a latent variable δ_R to connect with the auxiliary task module (see the next section). During training, the policy is guided by a reward function that encourages goal-reaching behaviors. Classic PointNav reward function [52] is used at each timestep t:

$$R^t_{\text{pointnav}} = -\beta_d \Delta_d - r_{\text{slack}} + \beta_{\text{succ}} \cdot I_{\text{succ}} \tag{2}$$

where Δ_d is the change in geodesic distance to the goal, r_{slack} is a step penalty to prevent unnecessary actions, t denotes the current time step, and I_{succ} is an indicator for successful navigation, β_d and β_{succ} are weight terms.

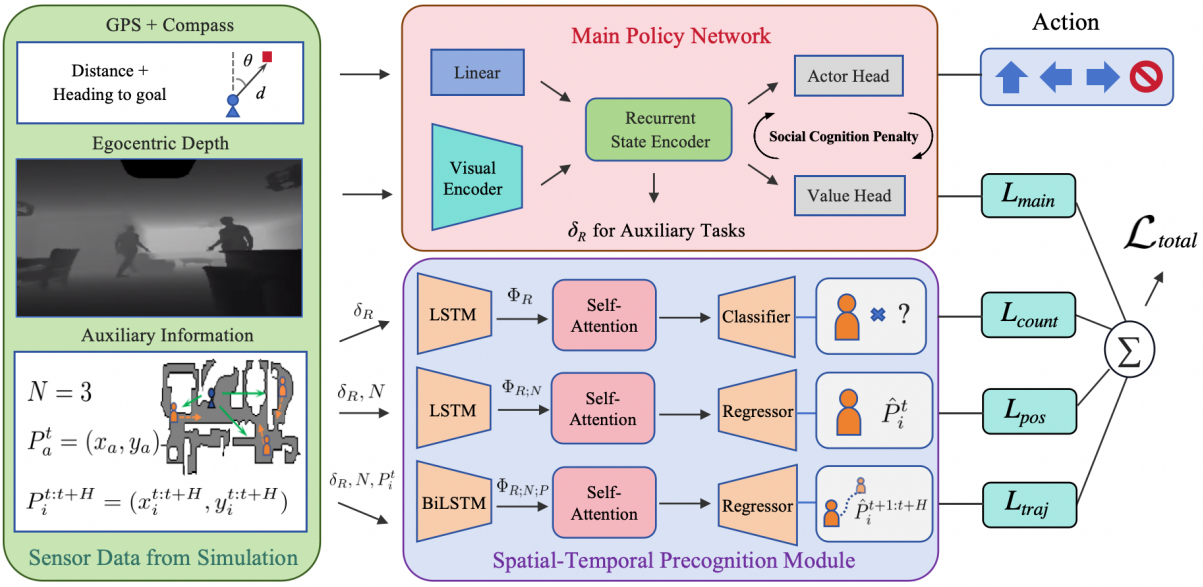


Fig. 2: ***Falcon* Overview**: The main policy network (top-right) takes Depth and GPS+Compass data as input. Its behavior is guided by the **Social Cognition Penalty**, which encourages socially compliant navigation and generates the main loss. During training, the output of the network's state encoder, combined with auxiliary information from the Habitat simulator, is processed by the **Spatial-Temporal Precognition Module** (bottom-right). Three socially-aware auxiliary tasks are then performed, producing auxiliary losses. The total loss is computed by weighting the main loss with the auxiliary losses.

However, PointNav's reward function is inadequate for SocialNav as it ignores dynamic environments and social interactions. To address this, we introduce **Social Cognition Penalty (SCP)**, a set of penalties designed to promote adherence to social norms. Specifically, it contains:

Obstacle Collision Penalty. Collisions with static obstacles or humans are penalized by

$$r_{\text{coll}} = \beta_s \cdot I_{\text{s\_coll}} + \beta_h \cdot I_{\text{h\_coll}} \tag{3}$$

where $I_{\text{s\_coll}}$ and $I_{\text{h\_coll}}$ are indicator variables representing collision with static obstacles and humans, respectively, and β_s and β_h are the corresponding penalty weights.

Human Proximity Penalty. This penalty ensures the agent maintains a safe distance from humans and is given by

$$r_{\text{prox}} = \sum_{i=1}^{N} \begin{cases} \beta_{\text{prox}} \cdot \exp\left(-d_i^t\right) & \text{if } d_i^t < 2.0 \text{ m} \\ 0 & \text{if } d_i^t \geq 2.0 \text{ m} \end{cases} \tag{4}$$

here $d_i^t = \|\tau_a(t) - \tau_i(t)\|$ is robot's distance to the i-th human at time step t. As d_i decreases, the penalty increases exponentially to prompt the agent to avoid humans. The penalty is removed when the robot is within 2.0 m to goal.

Trajectory Obstruction Penalty. This penalty discourages the robot from obstructing future H-step human trajectories. It anticipates potential obstructions by considering both current and future positions, with earlier overlaps penalized more heavily. This penalty is calculated as:

$$r_{\text{traj}} = \sum_{k=t+1}^{t+H} \sum_{i=1}^{N} \begin{cases} \beta_{\text{traj}} \cdot \left(\frac{1}{k-t+1}\right) & \text{if } d_{\text{traj\_}i}^k < 0.05 \text{ m} \\ 0 & \text{if } d_{\text{traj\_}i}^k \geq 0.05 \text{ m} \end{cases} \tag{5}$$

where $d_{\text{traj\_}i}^k = \|\tau_a(k) - \tilde{\tau}_i(k)\|$ is the distance between the robot's and human's future trajectories at the k-th time step. This penalty is also canceled when the robot nears its goal within 2.0 m.

The overall reward function is the combination of the goal-directed reward R_{pointnav}^t and the SCP penalties R_{scp}^t:

$$R_{\text{socialnav}}^t = R_{\text{pointnav}}^t - R_{\text{scp}}^t \tag{6}$$

where R_{scp}^t is defined as:

$$R_{\text{scp}}^t = r_{\text{coll}} + r_{\text{prox}} + r_{\text{traj}} \tag{7}$$

The main policy network is trained using DD-PPO [52] with the PPO loss $\mathcal{L}_{\text{main}}$.

C. Spatial-Temporal Precognition Module

This module utilizes three socially-aware auxiliary tasks to boost the robot's grasp of spatial-temporal dynamics. As shown in Fig. 2, similar networks with an LSTM encoder and a self-attention block are used for each auxiliary task. We discuss each of these tasks as follows.

Human Count Estimation. This task aims to estimate the overall number of humans and the output is a discrete value between 0 and M ($M = 6$ in our experiments). The LSTM takes as input the latent variable δ_R from the main policy network, and outputs an encoded Φ_R, which is then used for the self-attention layer [53] with $Q = K = V = \Phi_R$. The classifier ϕ_{count} then predicts the probabilities of human counts $\hat{n}_k$, using the attention output $\mathbf{A}^t$ at time step t :

$$\hat{n}_k = \phi_{\text{count}}(\mathbf{A}^t) \quad \text{for } k \in \{0, 1, \ldots, M\} \tag{8}$$

The loss is computed using Cross-Entropy:

$$\mathcal{L}_{\text{count}} = -\sum_{k=0}^{M} n_k \log(\hat{n}_k) \tag{9}$$

where n_k is the indicator for the true count, and $\hat{n}_k$ is the predicted probability for k humans.

Current Position Tracking. This task tracks the 2D locations of humans relative to the robot. Inputs are δ_R and the oracle number of human agents N in the scene. The LSTM output $\Phi_{R;N}$ is used as $Q = K = V = \Phi_{R;N}$ in the self-attention layer. The regressor ϕ_{pos} predicts each human's position $\hat{\mathbf{P}}_i^t$ at time t:

$$\hat{\mathbf{P}}_i^t = \phi_{\text{pos}}(\mathbf{A}^t) \tag{10}$$

We use Mean Squared Error (MSE) between the predicted and true positions, with a mask $\mathcal{M}$ to handle when there are less than M humans in the scene:

$$\mathcal{L}_{\text{pos}} = \frac{1}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} \|\hat{\mathbf{P}}_i^t - \mathbf{P}_i^t\|^2 \tag{11}$$

Future Trajectory Forecasting. This task predicts human trajectories over multiple time steps. It uses a bi-directional stacked LSTM due to its complexity. Inputs include δ_R, N, and current human positions P_i^t. The LSTM output $\Phi_{R;N;P}$ is used as $Q = K = V = \Phi_{R;N;P}$ in the attention layer. The regressor ϕ_{traj} predicts future trajectories over H steps:

$$\hat{\mathbf{P}}_i^{t+1:t+H} = \phi_{\text{traj}}(\mathbf{A}^t). \tag{12}$$

Similarly, MSE is applied with the mask $\mathcal{M}$:

$$\mathcal{L}_{\text{traj}} = \frac{1}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} \|\hat{\mathbf{P}}_i^{t+1:t+H} - \mathbf{P}_i^{t+1:t+H}\|^2 \tag{13}$$

The auxiliary loss is defined as

$$\mathcal{L}_{\text{aux}} = \mathcal{L}_{\text{count}} + \mathcal{L}_{\text{pos}} + \mathcal{L}_{\text{traj}} \tag{14}$$

During training, the main policy network and these auxiliary tasks are optimized together. The total loss function is a weighted sum of the main policy loss $\mathcal{L}_{\text{main}}$ and the auxiliary loss $\mathcal{L}_{\text{aux}}$:

$$\mathcal{L}_{\text{total}} = \beta_{\text{main}}\mathcal{L}_{\text{main}} + \beta_{\text{aux}}\mathcal{L}_{\text{aux}} \tag{15}$$

where β_{main} and β_{aux} are their respective loss weights.

IV. EXPERIMENTS

In this section, we first introduce the proposed benchmark and its two datasets, followed by the experiment setup and evaluation metrics. Next, we show the results of *Falcon* against prior methods. Finally, we conduct ablation analysis of two key components, Social Cognition Penalty and Spatial-Temporal Precognition Module.

A. Datasets

While existing datasets for SocialNav offer valuable environments, many of them [54], [55] lack diversity in scene types. Other datasets [34], [18], though rich in scenes, do not adequately balance human density and the human movement patterns are unnatural or do not have animation at all. To address these limitations, we introduce a benchmark with **Social-HM3D** and **Social-MP3D**, two simulation datasets derived from photo-realistic HM3D [56] and MP3D [57]. The statistical properties of our benchmark, in comparison with existing datasets, are shown in Table I. Our datasets offer a wide variety of environments with carefully calibrated human density, incorporating realistic human motions and natural movement patterns. These features ensure balanced interaction dynamics across diverse scenes, facilitating the development of more effective social navigation algorithms. We describe these key designs of our benchmark below:

Dataset	Num. Scenes	Scene Type	Max Num. Humans	Natural Motions
iGibson-SN [54]	15	residence	3	✗
Isaac Sim [55]	7	residence, office, depot, etc.	7	✓
HabiCrowd [34]	480	residence, office, gym, etc.	40	✗
HM3D-S [18]	900	residence, office, shop, etc.	3	✗
Social-HM3D	844	residence, office, shop, etc.	6	✓
Social-MP3D	72	residence, office, gym, etc.	6	✓

TABLE I: Statistics Comparison of SocialNav Datasets/Simulators: Our proposed Social-HM3D and Social-MP3D datasets feature extensive scene diversity and realistic interaction design, addressing the shortcomings of previous datasets which often relied on over-simplified human behaviors and imbalanced interaction dynamics.

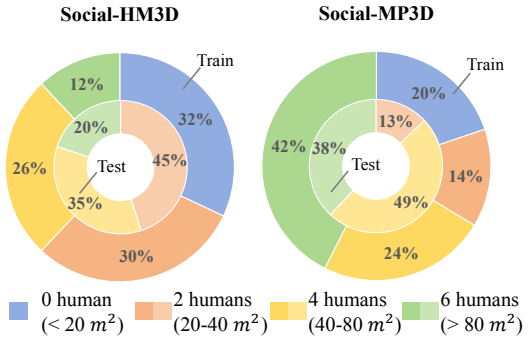


Fig. 3: Human Distribution by Scene Area in Social-HM3D and Social-MP3D (Train/Test): Our benchmark balances social density for human-robot interactions while avoiding overcrowding.

Realistic Agent Trajectories. Unlike previous datasets that leverage random walks or repetitive movements [54], [34], our dataset provides more realistic and reasonable human trajectories with goals. Each human is assigned two goals within the robot's navigation area, intended to encourage broader movement across the environment and increase the chance to interact with the robot.

Natural Human Behaviors. Our dataset realistically simulates human dynamics, with humans alternating between movement and rest, stopping once their tasks are complete. Walking speeds are randomized between 0.8 to 1.2 times the robot's speed, reflecting natural human movement. Humans react and avoid each other using ORCA algorithm [23]. Realistic walking animations for both humans and robots enhance the visual quality and authenticity of social interactions.

Reasonable Human Density. Our datasets group humans by scene area size to balance social interaction density. Fig. 3 shows the distribution of human numbers in the datasets,

which are carefully selected to ensure meaningful human-robot interactions while avoiding overcrowding that could hinder the movement of both the robot and humans.

Diversity and Scalability As shown in Table I, our datasets contain a broad range of environments, featuring 844 scenes derived from Social-HM3D and 72 scenes in Social-MP3D. The diversity of scenes and the inclusion of active humans at carefully chosen densities enhance its scalability to various tasks. Besides supporting SocialNav, our benchmark can be directly extended to social object/image navigation.

B. Experiment Setup

Metrics. Our benchmark metrics build upon existing works [58], [54] and focus on two principal perspectives: task completion and adherence to SocialNav objectives. For task completion, we use Success Rate (Suc.), Success weighted by Path Length (SPL) and Success weighted by Time Length (STL). For social norms, we evaluate Human-Robot Collision Rate (H-coll) and Personal Space Compliance (PSC). Considering the human collision radius is 0.3m and the robot is 0.25m, the PSC distance threshold is set to 1.0m.

Baseline Models. We mainly compare our method with a recent state-of-the-art social navigation method, namely Proximity-Aware [18]. It introduces two auxiliary tasks that model both the distance and direction of humans, effectively capturing their proximity at current. We also implement two classic baseline methods, including A-star (A*)[59] and ORCA[23] for comparison. A* and ORCA are both rule-based path-planning algorithms. ORCA has oracle access to agent positions and velocities to dynamically adjust the planned route. Both Proximity-Aware and our *Falcon* are RL-based. All methods only take depth image as visual inputs to ensure a fair comparison.

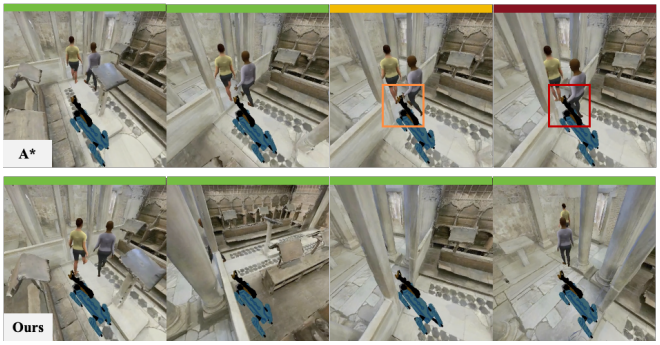
Implementation Details. We train the RL agents using DD-PPO algorithm [52] with identical hyperparameters. Each algorithm is run three times with different random seeds, and we report the mean and standard deviation of each metric. Our model is initialized with pre-trained weights from a PointNav model [56] and is fine-tuned for 10 million steps on the SocialNav task. Training is conducted on 4 Nvidia RTX 3090 GPUs with 8 parallel environments. Models are trained on the Social-HM3D train set and tested on both Social-HM3D and Social-MP3D, with Social-MP3D results used to evaluate *Falcon*'s zero-shot generalization.

C. Result Analysis

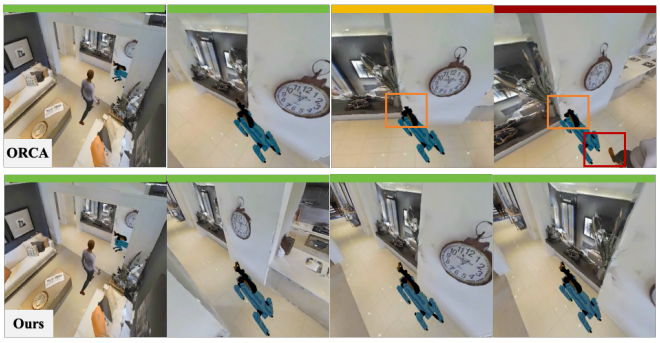
We conduct experiments to investigate several key aspects:

- Effectiveness of our algorithm against prior methods.
- Impacts of our auxiliary tasks.
- Individual and cooperative effects of the Social Cognition Penalty (SCP) and Spatial-Temporal Precognition Module (SPM).

Table II presents the comparison results on the Social-HM3D test set and zero-shot results on Social-MP3D test set. As we can see, the superior results demonstrate the efficiency of *Falcon* in both goal-reaching and social compliance, while the comparable zero-shot results highlight its generalization



(a) Person Following: A* algorithm does not wait for humans and causes a collision, while our method successfully follows.



(b) Intersection Encounter: ORCA algorithm fails to avoid a person and collides with a static obstacle, instead our method passes safely.

(c) Frontal Approach: The Proximity-Aware method collides when crossing in front directly, while our method avoids safely by anticipating human path.

Fig. 4: Comparisons of SocialNav Algorithms in Different Encounters: Our method outperforms other algorithms across various encounters. Green indicates safe behaviors, orange indicates risky behaviors (e.g., proximity to humans or collisions with obstacles), and red indicates unsafe behaviors (i.e., collisions with humans).

to unseen environments. Qualitatively, Fig. 4 provides examples illustrating various classic encounters in SocialNav tasks, where our method outperforms the others. Based on these results, we derive the following findings:

Finding 1: Future-aware methods are more efficient and safer than static and situation-aware approaches. Static path-planning algorithms like A* determine a fixed route and cannot adapt to dynamic environments, causing collision with humans (See Fig.4(a)). In contrast, situation-aware obstacle-avoidance approaches, such as ORCA and Proximity-Aware, react to the current environment by adjusting paths to avoid humans and obstacles. However, these

Dataset		Method	Suc. ↑	SPL ↑	STL ↑	PSC ↑	H-Coll ↓
Social-HM3D	Rule-Based	A* [59]	46.14±0.7	46.14±0.7	46.12±0.7	90.56±0.2	53.50±0.9
		ORCA [23]	38.91±0.1	38.91±0.1	38.44±0.1	90.55±0.4	47.52±1.7
	RL	Proximity-Aware [18]	20.11±1.3	18.57±1.9	19.51±1.5	**92.91**±0.5	**33.99**±0.7
		Falcon	**55.15**±0.6	**55.15**±0.7	**54.94**±0.7	89.56±1.4	42.96±1.1
Social-MP3D	Rule-Based	A* [59]	43.85±0.3	43.85±0.3	43.85±0.3	86.74±3.4	57.94±1.5
		ORCA [23]	40.38±0.3	40.38±0.3	39.51±0.2	91.76±0.4	47.16±0.2
	RL	Proximity-Aware [18]	18.45±1.4	17.09±2.8	16.41±1.5	**93.37**±0.9	**32.18**±3.3
		Falcon	**55.05**±0.7	**55.04**±0.6	**54.80**±1.0	90.01±1.2	42.19±0.9

TABLE II: Performance Evaluation of SocialNav Tasks for Rule-Based and RL-Based Methods on Social-HM3D (upper group) and Social-MP3D (lower group). Data in the table represents percentages. We **bold** the best results and underline the second best results.

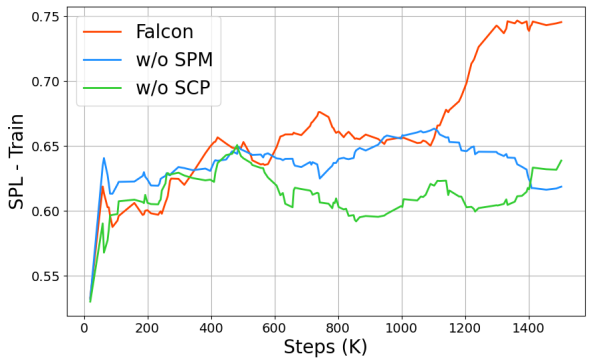


Fig. 5: Training Curve of SPL for Ablation Study: The full *Falcon* model with SPM and SCP converges faster and performs better.

SPM. Count	SPM. Pos	SPM. Traj	SCP	Suc. ↑	SPL ↑	STL ↑	PSC ↑	H-Coll ↓
PointNav (w/o Aux. Task)				40.94	34.14	11.50	**90.82**	53.54
✓				51.43	51.42	51.16	90.53	46.46
	✓			53.17	53.17	52.95	90.06	44.07
		✓		54.00	53.99	53.92	89.46	43.88
			✓	51.24	51.24	51.08	90.41	48.11
✓	✓	✓		53.63	53.63	53.40	89.33	44.89
✓	✓	✓	✓	**55.15**	**55.15**	**54.94**	89.56	**42.96**

TABLE III: Ablation Study for *Falcon*. The model trained solely with the PointNav algorithm [56] serves as the baseline. SPM.Count, SPM.Pos, and SPM.Traj refer to three auxiliary tasks: Humanoid Count Estimation, Current Position Tracking, and Future Trajectory Forecasting. Data in the table are percentages.

methods have limitations: rerouting often takes time, and delayed reactions increase collision risks. ORCA assumes unrestricted movement and causes a collision with static obstacles (See Fig.4(b)). Similarly, the Proximity-Aware method struggles due to its inability to predict human movement, leading to failed short-term adjustments (See Fig.4(c)). *Falcon* is able to proactively adjust to dynamic human movements and reach the goal location effectively.

Finding 2: Auxiliary tasks contribute to performance improvement, with trajectory prediction playing the most significant role. Table III presents experiments with various auxiliary task choices. Each task individually improves navigation performance compared to the PointNav baseline [56]. Among them, trajectory forecasting (SPM.Traj) proves to be the most effective auxiliary task, with the success rate notably increasing from 40.94% to 54.00%. This highlights the value of explicit trajectory prediction in SocialNav.

Finding 3: SCP coordinates and complements SPM, leading to a significant improvement and faster training. Table III shows that SCP plays a critical role in enhancing the model's performance, particularly when integrated with the SPM. As we can see, the SPM with its three auxiliary tasks (Count, Pos, and Traj) does not improve much upon training them individually (SPM's 53.63% vs. SPM.Pos's 53.17% and SPM.Traj's 54.00%). With SCP, the full system achieves significant improvement over SPM (55.15% vs. 53.63%). Also, as shown in Fig. 5, the model trained with both SPM and SCP exhibits faster convergence during training (evident before 1400K steps). These results show that SPM tasks cannot be effectively integrated without the guidance of SCP, which helps the model balance tasks and better leverage available information.

V. CONCLUSIONS

In this paper, we introduce a novel SocialNav benchmark with two datasets, Social-HM3D and Social-MP3D. We propose *Falcon*, a future-aware method for social navigation in realistic human-populated scenes. Our method achieves state-of-the-art results over rule-based and recent reinforcement learning models on our proposed new benchmark. We believe our benchmark, with our method, will facilitate future research and applications on social navigation. The superior performance in success rates and collision avoidance highlights the effectiveness of combining trajectory prediction into navigation policy.

VI. ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 62306257), and the Guangzhou Municipal Science and Technology Project (No. 2024A03J0619). This work was also supported by the Meituan Academy of Robotics Shenzhen, the Guangzhou-HKUST(GZ) Joint Funding Program (Grant No.2023A03J0008) and Education Bureau of Guangzhou Municipality.

REFERENCES

[1] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh, "Core challenges of social robot navigation: A survey," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 3, pp. 1–39, 2023.

[2] A. Kapoor, S. Swamy, P. Bachiller, and L. J. Manso, "Socnavgym: a reinforcement learning gym for social navigation," in *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2023, pp. 2010–2017.

[3] W. Wang, L. Mao, R. Wang, and B.-C. Min, "Multi-robot cooperative socially-aware navigation using multi-agent reinforcement learning," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 12 353–12 360.

[4] N. Hirose, D. Shah, K. Stachowicz, A. Sridhar, and S. Levine, "Selfi: Autonomous self-improvement with reinforcement learning for social navigation," *arXiv preprint arXiv:2403.00991*, 2024.

[5] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 6015–6022.

[6] K. Li, Y. Xu, J. Wang, and M. Q.-H. Meng, "Sarl: Deep reinforcement learning based human-aware navigation for mobile robot in indoor environments," in *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2019, pp. 688–694.

[7] C. Pérez-D'Arpino, C. Liu, P. Goebel, R. Martín-Martín, and S. Savarese, "Robot navigation in constrained pedestrian environments using reinforcement learning," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1140–1146.

[8] J. Liang, L. Jiang, J. C. Niebles, A. G. Hauptmann, and L. Fei-Fei, "Peeking into the future: Predicting future person activities and locations in videos," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 5725–5734.

[9] H. Nishimura, B. Ivanovic, A. Gaidon, M. Pavone, and M. Schwager, "Risk-sensitive sequential action control with multi-modal human trajectory forecasting for safe crowd-robot interaction," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 11 205–11 212.

[10] K. Mangalam, Y. An, H. Girase, and J. Malik, "From goals, waypoints & paths to long term human trajectory forecasting," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 15 233–15 242.

[11] J. Liang, L. Jiang, K. Murphy, T. Yu, and A. Hauptmann, "The garden of forking paths: Towards multi-future trajectory prediction," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 508–10 518.

[12] J. Liang, L. Jiang, and A. Hauptmann, "Simaug: Learning robust representations from simulation for trajectory prediction," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16*. Springer, 2020, pp. 275–292.

[13] H. Zhao, J. Gao, T. Lan, C. Sun, B. Sapp, B. Varadarajan, Y. Shen, Y. Shen, Y. Chai, C. Schmid, *et al.*, "Tnt: Target-driven trajectory prediction," in *Conference on Robot Learning*. PMLR, 2021, pp. 895–904.

[14] A. Francis, C. Pérez-d'Arpino, C. Li, F. Xia, A. Alahi, R. Alami, A. Bera, A. Biswas, J. Biswas, R. Chandra, *et al.*, "Principles and guidelines for evaluating social robot navigation algorithms," *arXiv preprint arXiv:2306.16740*, 2023.

[15] S. Liu, P. Chang, W. Liang, N. Chakraborty, and K. Driggs-Campbell, "Decentralized structural-rnn for robot crowd navigation with deep reinforcement learning," in *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021, pp. 3517–3524.

[16] P. Kothari, S. Kreiss, and A. Alahi, "Human trajectory forecasting in crowds: A deep learning perspective," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7386–7400, 2022.

[17] S. Holk, D. Marta, and I. Leite, "Polite: Preferences combined with highlights in reinforcement learning," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 2288–2295.

[18] E. Cancelli, T. Campari, L. Serafini, A. X. Chang, and L. Ballan, "Exploiting proximity-aware tasks for embodied social navigation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 10 957–10 967.

[19] F. Xia, W. B. Shen, C. Li, P. Kasimbeg, M. E. Tchapmi, A. Toshev, R. Martín-Martín, and S. Savarese, "Interactive gibson benchmark: A benchmark for interactive navigation in cluttered environments," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 713–720, 2020.

[20] X. Puig, E. Undersander, A. Szot, M. D. Cote, T.-Y. Yang, R. Partsey, R. Desai, A. W. Clegg, M. Hlavac, S. Y. Min, *et al.*, "Habitat 3.0: A co-habitat for humans, avatars and robots," *arXiv preprint arXiv:2310.13724*, 2023.

[21] P. T. Singamaneni, P. Bachiller-Burgos, L. J. Manso, A. Garrell, A. Sanfeliu, A. Spalanzani, and R. Alami, "A survey on socially aware robot navigation: Taxonomy and future challenges," *The International Journal of Robotics Research*, p. 02783649241230562, 2024.

[22] R. Möller, A. Furnari, S. Battiato, A. Härmä, and G. M. Farinella, "A survey on human-aware robot navigation," *Robotics and Autonomous Systems*, vol. 145, p. 103837, 2021.

[23] J. Van Den Berg, S. J. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," in *Robotics Research: The 14th International Symposium ISRR*. Springer, 2011, pp. 3–19.

[24] J. Van den Berg, M. Lin, and D. Manocha, "Reciprocal velocity obstacles for real-time multi-agent navigation," 2008.

[25] P. Long, W. Liu, and J. Pan, "Deep-learned collision avoidance policy for distributed multiagent navigation," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 656–663, 2017.

[26] G. S. Aoude, B. D. Luders, J. M. Joseph, N. Roy, and J. P. How, "Probabilistically safe motion planning to avoid dynamic obstacles with uncertain motion patterns," *Autonomous Robots*, vol. 35, no. 1, pp. 51–76, 2013.

[27] J. Guzzi, A. Giusti, L. M. Gambardella, G. Theraulaz, and G. A. Di Caro, "Human-friendly robot navigation in dynamic environments," in *IEEE international conference on robotics and automation*. IEEE, 2013, pp. 423–430.

[28] G. Ferrer, A. Garrell, and A. Sanfeliu, "Social-aware robot navigation in urban environments," in *Proc. of thr European Conference on Mobile Robots*, 2013.

[29] Y. F. Chen, M. Everett, M. Liu, and J. P. How, "Socially aware motion planning with deep reinforcement learning," 2017.

[30] Y. Lu, X. Ruan, and J. Huang, "Deep reinforcement learning based on social spatial–temporal graph convolution network for crowd navigation," *Machines*, vol. 10, no. 8, p. 703, 2022.

[31] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," 2019.

[32] A. Rudenko, T. P. Kucner, C. S. Swaminathan, R. T. Chadalavada, K. O. Arras, and A. J. Lilienthal, "Thör: Human-robot navigation data collection and accurate motion trajectories dataset," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 676–682, 2020.

[33] R. Martin-Martin, M. Patel, H. Rezatofighi, A. Shenoi, J. Gwak, E. Frankel, A. Sadeghian, and S. Savarese, "JRDB: A dataset and benchmark of egocentric robot visual perception of humans in built environments," *IEEE transactions on pattern analysis and machine intelligence*, 2021.

[34] A. D. Vuong, T. T. Nguyen, M. N. VU, B. Huang, D. Nguyen, H. T. T. Binh, T. Vo, and A. Nguyen, "Habicrowd: A high performance simulator for crowd-aware visual navigation," *arXiv preprint arXiv:2306.11377*, 2023.

[35] P. Mirowski, R. Pascanu, F. Viola, H. Soyer, A. J. Ballard, A. Banino, M. Denil, R. Goroshin, L. Sifre, K. Kavukcuoglu, *et al.*, "Learning to navigate in complex environments," *arXiv preprint arXiv:1611.03673*, 2016.

[36] M. Jaderberg, V. Mnih, W. M. Czarnecki, T. Schaul, J. Z. Leibo, D. Silver, and K. Kavukcuoglu, "Reinforcement learning with unsupervised auxiliary tasks," *arXiv preprint arXiv:1611.05397*, 2016.

[37] X. Lin, H. Baweja, G. Kantor, and D. Held, "Adaptive auxiliary task weighting for reinforcement learning," *Advances in neural information processing systems*, vol. 32, 2019.

[38] J. Ye, D. Batra, E. Wijmans, and A. Das, "Auxiliary tasks speed up learning point goal navigation," in *Conference on Robot Learning*. PMLR, 2021, pp. 498–516.

[39] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *The International Journal of Robotics Research*, vol. 39, no. 8, pp. 895–935, 2020.

[40] R. Huang, H. Xue, M. Pagnucco, F. Salim, and Y. Song, "Multimodal trajectory prediction: A survey," *arXiv preprint arXiv:2302.10463*, 2023.

[41] A. Elnagar, "Prediction of moving objects in dynamic environments using kalman filters," in *Proceedings 2001 IEEE International Sympo-*

sium on Computational Intelligence in Robotics and Automation (Cat. No. 01EX515). IEEE, 2001, pp. 414–419.

[42] S. Zernetsch, S. Kohnen, M. Goldhammer, K. Doll, and B. Sick, "Trajectory prediction of cyclists using a physical model and an artificial neural network," in *2016 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2016, pp. 833–838.

[43] P. Coscia, F. Castaldo, F. A. Palmieri, A. Alahi, S. Savarese, and L. Ballan, "Long-term path prediction in urban scenarios using circular distributions," *Image and Vision Computing*, vol. 69, pp. 81–91, 2018.

[44] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social lstm: Human trajectory prediction in crowded spaces," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 961–971.

[45] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts, and A. J. Lilienthal, "Enabling flow awareness for mobile robots in partially observable environments," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1093–1100, 2017.

[46] A. Vemula, K. Muelling, and J. Oh, "Modeling cooperative navigation in dense human crowds," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 1685–1692.

[47] D. Vasquez, "Novel planning-based algorithms for human motion prediction," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 3317–3322.

[48] A. Rudenko, L. Palmieri, and K. O. Arras, "Predictive planning for a mobile robot in human environments," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA), Works. on AI Planning and Robotics*, 2017.

[49] E. Rehder, F. Wirth, M. Lauer, and C. Stiller, "Pedestrian prediction by planning using deep neural networks," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 5903–5908.

[50] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[51] A. Graves and A. Graves, "Long short-term memory," *Supervised sequence labelling with recurrent neural networks*, pp. 37–45, 2012.

[52] E. Wijmans, A. Kadian, A. Morcos, S. Lee, I. Essa, D. Parikh, M. Savva, and D. Batra, "Dd-ppo: Learning near-perfect pointgoal navigators from 2.5 billion frames," *arXiv preprint arXiv:1911.00357*, 2019.

[53] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.

[54] C. Li, J. Jang, F. Xia, R. Martín-Martín, C. D'Arpino, A. Toshev, A. Francis, E. Lee, and S. Savarese, "igibson challenge 2021," Online, 2021, [Online; accessed: 2024-08-23]. [Online]. Available: http://svl.stanford.edu/igibson/challenge.html

[55] V. Makoviychuk, L. Wawrzyniak, Y. Guo, M. Lu, K. Storey, M. Macklin, D. Hoeller, N. Rudin, A. Allshire, A. Handa, *et al.*, "Isaac gym: High performance gpu-based physics simulation for robot learning," *arXiv preprint arXiv:2108.10470*, 2021.

[56] S. K. Ramakrishnan, A. Gokaslan, E. Wijmans, O. Maksymets, A. Clegg, J. Turner, E. Undersander, W. Galuba, A. Westbury, A. X. Chang, *et al.*, "Habitat-matterport 3d dataset (hm3d): 1000 large-scale 3d environments for embodied ai," *arXiv preprint arXiv:2109.08238*, 2021.

[57] A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Niessner, M. Savva, S. Song, A. Zeng, and Y. Zhang, "Matterport3d: Learning from rgb-d data in indoor environments," *arXiv preprint arXiv:1709.06158*, 2017.

[58] P. Anderson, A. Chang, D. S. Chaplot, A. Dosovitskiy, S. Gupta, V. Koltun, J. Kosecka, J. Malik, R. Mottaghi, M. Savva, *et al.*, "On evaluation of embodied navigation agents," *arXiv preprint arXiv:1807.06757*, 2018.

[59] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.