# Following the Human Thread in Social Navigation

Luca Scofano[1*] Alessio Sampieri[1*] Tommaso Campari[2*] Valentino Sacco[1]

Indro Spinelli[1] Lamberto Ballan[3] Fabio Galasso[1]

[1] Sapienza University of Rome [2] Fondazione Bruno Kessler (FBK) [3] University of Padova

## Abstract

*The paper proposes a Social Dynamics Adaptation model (SDA) for Social Navigation, which involves a robot's ability to navigate human-centric environments while maintaining a safe distance and adhering to social norms. The key challenge is to process human trajectories, which are partially observable from the robot's perspective and complex to compute. The proposed SDA model uses a two-stage Reinforcement Learning framework: the first stage involves learning to encode human trajectories and the second stage infers social dynamics from the robot's state-action history. This approach has been tested on the Habitat 3.0 platform, achieving state-of-the-art performance in finding and following humans. The extended version of this work is available at: https://arxiv.org/abs/2404.11327.*

## 1. Introduction

Embodied Artificial Intelligence (EAI) has significantly advanced traditional navigation techniques by introducing robots into real-life environments. Now, when navigating human-centric environments, the agents must consider human movements and behaviors, changing the focus to socially aware navigation. Social navigation agents need to locate, track, and interact with humans safely and in a socially acceptable manner.

Previous studies have characterized Social Navigation as a variation of PointGoal Navigation, where agents aim to reach destinations while considering human movements as dynamic obstacles [3, 5, 6]. Habitat 3.0 [4] introduced a lifelike environment with human avatars, where human-agent interactions are present in a dynamic, controlled, and safe setting. However, this complexity also presents challenges other than collision avoidance, such as locating/following the humans in the scene.

Existing methods for Social Navigation often rely on privileged information unavailable in real-world scenarios or fail to capture social dynamics and norms adequately. For instance, [5, 6] and the current SoA model on Habitat 3.0 [4] use GPS and compass sensors for perfect humanoid localization, which may be impractical during real-world inference. In contrast, [3, 7] do not consider social factors influencing human behaviour. [1] models some social factors but falls short in capturing the cooperative nature of social agents.

This paper introduces a novel Social Dynamics Adaptation model (SDA) to address these limitations. SDA is a two-stage model (Fig. 1): the first stage trains a base policy using human trajectories encoded into a latent vector, and the second stage infers social dynamics from the robot's state and action history. Unlike prior methods, deploying SDA after this last stage allows the agent to react to human movements without having access to privileged information.

## 2. Methodology

**Stage 1: Social Policy** In the first stage, SDA takes as input image features $x_t$ and the action at the previous timestep $a_{t-1}$. We add another input to this pipeline, namely a latent vector $z_t$ built by encoding the humanoid privileged information, $e_{t-N:t-1}$ as $z_t = \mu(e_{t-N:t-1})$ and thus we can define the next action $a_t = \pi(x_t, a_{t-1}, z_t)$.

Where the trajectory encoder ($\mu$) is implemented as Multilayer Perceptrons (MLPs). Intuitively, training everything with the same objective $z_t$ encodes the social dynamics that led the agent to maximize its reward, adapting to human movement patterns. The objective retains its usual formulation [5] without any explicit reference to the human trajectories.

**Stage 2: Social Dynamics Regression** We aim to extract and exploit social cues directly from the robot's perception and eliminate the need for auxiliary devices like GPS track-

[*] The authors contributed equally to this work.

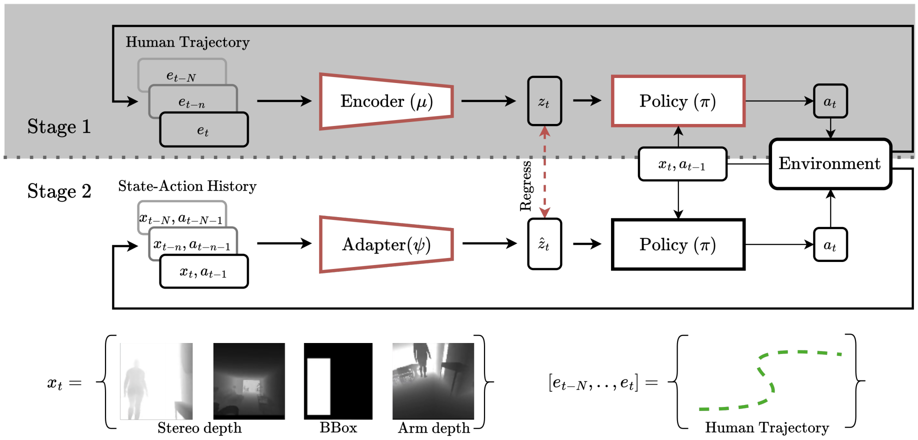


Figure 1. SDA framework. Stage 1, jointly learns to encode human trajectories and a motion policy. Stage 2 infers the social dynamics from the state-action history and feeds them to the frozen policy.

| Models | hGPS | traj. | S ↑ | SPS ↑ | F ↑ | CR ↓ | ES ↑ |
|---|---|---|---|---|---|---|---|
| Heuristic Expert [4] | - | - | 1.00 | 0.97 | 0.51 | 0.52 | - |
| Baseline [4] | GT | | $0.97^{\pm 0.00}$ | $0.65^{\pm 0.00}$ | $0.44^{\pm 0.01}$ | $0.51^{\pm 0.03}$ | $0.55^{\pm 0.01}$* |
| Baseline+Proximity [1][1] | GT | | $0.97^{\pm 0.01}$ | $0.64^{\pm 0.00}$ | $0.57^{\pm 0.01}$ | $0.58^{\pm 0.03}$ | $0.63^{\pm 0.02}$ |
| SDA - Stage1 | | GT | $0.92^{\pm 0.00}$ | $0.46^{\pm 0.01}$ | $0.44^{\pm 0.02}$ | $0.61^{\pm 0.02}$ | $0.50^{\pm 0.01}$ |
| Baseline [4] | | | $0.76^{\pm 0.02}$ | $0.34^{\pm 0.01}$ | $0.29^{\pm 0.01}$ | **0.48**$^{\pm 0.03}$ | $0.40^{\pm 0.02}$* |
| Baseline+Proximity [1] | | | $0.85^{\pm 0.02}$ | $0.41^{\pm 0.02}$ | $0.37^{\pm 0.01}$ | $0.58^{\pm 0.02}$ | $0.41^{\pm 0.01}$ |
| SDA - Stage2 | | | **0.91**$^{\pm 0.01}$ | **0.45**$^{\pm 0.01}$ | **0.39**$^{\pm 0.01}$ | $0.57^{\pm 0.02}$ | **0.43**$^{\pm 0.02}$ |

Table 1. Social Navigation results. GT denotes ground truth privileged information and * denotes reproduced results.

ers on humanoids. Inspired by [2], we introduce the "social dynamics" module (Adapter), parametrized by an MLP $\psi$ that takes as input the recent history of the robot's states $x_{t-N:t-1}$ and actions $a_{t-N:t-1}$ to generate a new latent vector $\hat{z}_t = \psi(x_{t-N:t-1}, a_{t-N:t-1})$.

We obtain the state-action history by deploying the agent in the environment with optimal policy $\pi^*$ obtained after the first stage and the latent vector $\hat{z}_t$ and get $a_t = \pi^*(x_t, a_{t-1}, \hat{z}_t)$. During this process, we optimize the Adapter, MLP, with a supervised regression objective, Mean Squared Error (MSE), to recover the original information in $z_t$.

Once we finalize the Adapter training, instead of relying on the privileged information, we can depend upon the robot's states $x_{t-N:t-1}$ and actions $a_{t-N:t-1}$ to generate $\hat{z}_t$, serving as an estimate of the actual latent social dynamics vector $z_t$. Doing so enables the agent to estimate social dynamics online, improving its performance in dynamic environments and enhancing its social navigation capabilities, freeing it from dependence upon external sensors.

## 3. Results

The first section of Table 1 shows the performance of a heuristic expert with extensive information. The following section sets different upper bounds using models trained and tested using ground truth (GT) privileged information. The third section reports methods tested without privileged information. SDA-Stage 1 provides a lower upper bound for S (finding success) and SPS (shortest path to human) due to the absence of the humanoid GPS (hGPS).

SDA-Stage2 maintains the performances of Stage 1 by correctly inferring social dynamics, despite the lack of human trajectory input. SDA-Stage2 outperforms others in finding related metrics, increasing $S$ and $SPS$ by $6\%$ and $4\%$. In fact, it locates the humanoid faster (438 vs. 540 avg. steps) on a maximum episode length of $1500$.

We also improve episode success (ES), where the agent finds and follows the humanoid for at least $400$ steps without colliding. We follow for longer (390 vs. 218 avg. steps), improving the following rate (F). Episodes do not necessarily terminate after $400$, thus our higher reported collision rate (CR). Otherwise, the $CR$ would be $0.39$ for SDA and $0.38$ for the best baseline.

**Acknowledgements** This project was supported by PNRR MUR project PE0000013-FAIR. We acknowledge the CINECA award under the ISCRA initiative for the availability of high-performance computing resources and support.

## References

[1] Enrico Cancelli, Tommaso Campari, Luciano Serafini, Angel X. Chang, and Lamberto Ballan. Exploiting proximity-aware tasks for embodied social navigation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10957–10967, 2023. 1, 2

[2] Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. RMA: Rapid motor adaptation for legged robots. In *Robotics: Science and Systems*, 2021. 2

[3] Ruslan Partsey, Erik Wijmans, Naoki Yokoyama, Oles Dobosevych, Dhruv Batra, and Oleksandr Maksymets. Is mapping necessary for realistic pointgoal navigation? In *CVPR*, 2022. 1

[4] Xavier Puig, Eric Undersander, Andrew Szot, Mikael Dallaire Cote, Tsung-Yen Yang, Ruslan Partsey, Ruta Desai, Alexander William Clegg, Michal Hlavac, So Yeon Min, Vladimír Vondruš, Theophile Gervet, Vincent-Pierre Berges, John M. Turner, Oleksandr Maksymets, Zsolt Kira, Mrinal Kalakrishnan, Jitendra Malik, Devendra Singh Chaplot, Unnat Jain, Dhruv Batra, Akshara Rai, and Roozbeh Mottaghi. Habitat 3.0: A co-habitat for humans, avatars and robots, 2023. 1, 2

[5] Erik Wijmans, Abhishek Kadian, Ari Morcos, Stefan Lee, Irfan Essa, Devi Parikh, Manolis Savva, and Dhruv Batra. DD-PPO: Learning near-perfect pointgoal navigators from 2.5 billion frames. In *ICLR*, 2019. 1

[6] Joel Ye, Dhruv Batra, Abhishek Das, and Erik Wijmans. Auxiliary Tasks and Exploration Enable ObjectNav. In *ICCV*, 2021. 1

[7] Naoki Yokoyama, Qian Luo, Dhruv Batra, and Sehoon Ha. Learning Robust Agents for Visual Navigation in Dynamic Environments: The Winning Entry of iGibson Challenge 2021. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022. 1