From Crowd Motion Prediction to Robot Navigation in Crowds*

Sriyash Poddar[1], Christoforos Mavrogiannis[2], Siddhartha S. Srinivasa[2]

***Abstract*— To navigate safely and efficiently within crowds, robots need models for crowd motion prediction. Building such models is hard due to the high dimensionality of multiagent domains and the challenge of collecting or simulating interaction-rich crowd-robot demonstrations. While there has been important progress on models for offline pedestrian motion forecasting, transferring their performance on real robots is nontrivial due to close interaction settings and novelty effects on users. In this paper, we investigate the utility of a recent state-of-the-art motion prediction model (S-GAN) for crowd navigation tasks. We incorporate this model into a model predictive controller (MPC) and deploy it on a self-balancing robot which we subject to a diverse range of crowd behaviors in the lab. We demonstrate that while S-GAN motion prediction accuracy transfers to the real world, its value is not reflected on navigation performance, measured with respect to safety and efficiency; in fact, the MPC performs indistinguishably even when using a simple constant-velocity prediction model, suggesting that substantial model improvements and user-centered optimization criteria might be needed to yield significant gains for crowd navigation tasks.**

I. Introduction

Large-scale deep learning methods [2, 5, 8, 20, 24, 25, 34] have dramatically improved the state-of-the-art in prediction accuracy across standard benchmarks [14, 21]. While these models have been the foundation of recent real-world robot demonstrations [1, 3, 4, 6, 10, 15], scaling their performance to complex environments like pedestrian domains, warehouses, or hospitals is challenging as these environments feature close interaction settings, large space of behavior, and limited rules.

To address these challenges, many approaches learn end-to-end deep learning models from simulated crowd-robot interactions [3, 4, 6, 15]. While typical crowd simulators [9, 30] produce realistic behaviors, some of their core assumptions limit their relevance to crowd-navigation tasks. For instance, Fraichard and Levesy [7] showed that the assumptions of omniscience and homogeneity of existing crowd simulators give rise to behaviors that would be unsafe to execute on a real robot. Further, Mavrogiannis et al. [19] showed that a non-reactive, non-collision-avoiding agent is safer than ORCA-simulated agents in an ORCA-simulated world [30] due to the overly submissive behaviors this model may exhibit.

[1]Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur, Kharagpur, India. Email: poddarsriyash@iitkgp.ac.in

[2]Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, USA. Email: {cmavro, siddh}@cs.washington.edu

*This paper is a short summary of our prior work [18, 22]. Footage from our experiments can be found here and our code implementation here.

This work was (partially) funded by the Honda Research Institute USA, the National Science Foundation NRI (#2132848) and CHS (#2007011), DARPA RACER (#HR0011-21-C-0171), the Office of Naval Research (#N00014-17-1-2617-P00004 and #2022-016-01 UW), and Amazon.

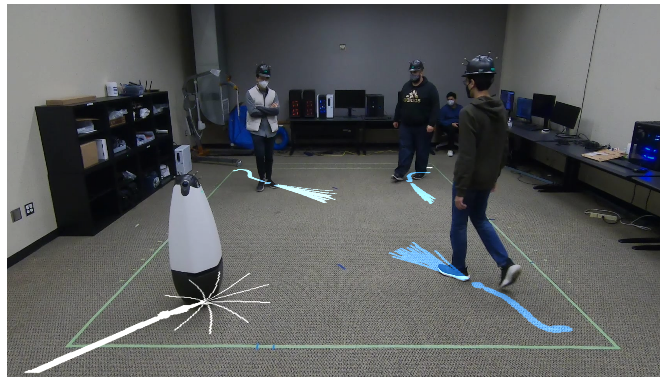


Fig. 1: Honda's experimental ballbot [11] navigates next to three users in our lab. Agents' past trajectories and distribution of future actions are shown. In this paper, we approach the question of how human motion prediction accuracy translates into robot navigation performance in crowded environments.

Other approaches learn models of human motion prediction from pedestrian datasets and incorporate them into navigation controllers [1, 4, 12, 13, 18, 26, 28, 29, 32, 35]. However, the pedestrian datasets commonly used [14, 21] feature well-structured, goal-directed, and cooperative motion. These settings represent a narrow subset of the behavior that a robot would encounter in the real world. This behavior is so prevalent in those datasets that according to Schöller et al. [27], even constant-velocity (CV) prediction, a very simple, analytical model, performs comparably to recent state-of-the-art (SOTA) deep models. Therefore, while the SOTA in human motion prediction keeps improving, it is unclear what its relevance is for robot navigation in crowds.

Inspired by these observations, we ask the question:

> *To what extent does crowd motion prediction accuracy translate to robot navigation performance in crowd navigation tasks?*

To approach this question, we investigate the transfer of a SOTA model (S-GAN [8]) from offline to onboard performance and its implications for crowd navigation. We integrate the S-GAN into a MPC and deploy it on a self-balancing robot (see Fig. 1), which we subject to diverse crowd conditions in the lab. Through extensive experiments, we find that while the onboard prediction accuracy of S-GAN is superior to a simple CV baseline, the MPC navigation performance (safety and efficiency) is indistinguishable, suggesting that substantial prediction improvements and new validated optimization criteria may be needed to improve navigation performance. Our evaluation includes novel benchmarking experiments involving a variety of crowd conditions that may serve as evaluation settings for future crowd navigation experiments.

II. Problem Statement

We consider a workspace $\mathcal{W} \subseteq \mathbb{R}^2$ where a robot navigates among n human agents. We denote by $s \in \mathcal{W}$ the robot state and by $s^i \in \mathcal{W}$ the agent state $i \in \mathcal{N} = \{1, \dots n\}$. The robot is navigating from a state s_0 towards a goal state g whereas agent $i \in \mathcal{N} = \{1, \dots n\}$ is navigating from s_0^i towards a destination g^i. The robot is unaware of g^i but we assume that it is fully observing the world state $(s_t, s_t^{1:n})$ at every timestep t. Maintaining a history of states for all agents, the robot predicts their future trajectories using a model f. In this paper, our goal is to investigate whether the prediction accuracy of f translates to robot navigation performance. As a proxy for navigation performance, we consider metrics capturing safety and efficiency of robot motion.

III. Human Motion Prediction

We treat human motion prediction as trajectory prediction over a horizon T given a past trajectory of horizon h.

A. Probabilistic Trajectory Prediction

We denote by $s^i_{t-h:t} \in \mathcal{W}^h$ the partial trajectory of an agent $i \in \mathcal{N}$ of horizon h and by $s^i_{t:t+T} \in \mathcal{W}^T$ the future trajectory until time T. Consider a joint state prediction model $f : \mathcal{W}^{n\times h} \to \mathcal{W}^{n\times T}$, which takes as input the joint states of the agents $\boldsymbol{s}^{1:n}_{t-h:t}$ and predicts the future states $\hat{\boldsymbol{s}}^{1:n}$.

$$f\left(s^1_{t-h:t}, \dots s^n_{t-h:t}\right) = \left(\hat{s}^1_{t:t+T}, \dots \hat{s}^n_{t:t+T}\right) = \hat{\boldsymbol{s}}^{1:n}$$

We denote the distribution of future states for an agent $i \in \mathcal{N}$ as $p(\hat{s}^i_{t:t+T})$, and the joint distribution of states is represented as $p(\hat{\boldsymbol{s}}^{1:n})$. The prediction model $f : \mathcal{W}^{t\times n} \times \mathcal{W}^{T\times n} \to [0,1]$ is a conditional distribution; denoting the distribution of the future trajectories given past trajectories of all the agents i.e f corresponds to $p(\hat{\boldsymbol{s}}^{1:n} | s^{1:n}_{t-h:t})$. In this paper, we adopt a probabilistic trajectory prediction mechanism f using Social GAN (S-GAN), a state-of-the-art model from Gupta et al. [8]. The generative S-GAN model is able to output samples for the future states $s^{1:n}_{t:t+T} \sim p(s^{1:n}_{t:t+T} | s^{1:n}_{0:t})$.

B. Offline Prediction Performance

Schöller et al. [27] compared the Average Displacement Error (ADE) and Final Displacement Error (FDE) of S-GAN-based prediction against CV prediction and CV prediction with added noise (CVN), showing that the latter ones perform comparably across the scenes in the ETH [21] and UCY [14] datasets. In Fig. 2, we compare their *multistep* prediction performance (i.e., the L2-norm between the predicted position and ground truth at each timestep of prediction), which is informative for navigation tasks. We note that the mixed performance of S-GAN (lower on Zara1, similar or outperformed on others) can be attributed to the datasets mostly consisting of linear segments well approximated by CV/CVN.

IV. MPC with Probabilistic Multiagent Trajectory Prediction

We integrate prediction models from Sec. III into an MPC for crowd navigation.

A. MPC for Navigation in Crowds

We employ a discrete MPC formulation for navigation in a multiagent environment:

$$\begin{aligned} \boldsymbol{u}^* = \arg\min_{\boldsymbol{u}\in\boldsymbol{\mathcal{U}}} & \mathcal{J}(\boldsymbol{s}, \hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}^{1:n}) \\ s.t.\ & s_{t+1} = g(s_t, u_t) \\ & (\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}^{1:n}) = f(s^{1:n}_{t-h:t}, s_{t-h:t}),\ i \in \mathcal{N} \end{aligned}, \tag{1}$$

where: $\boldsymbol{s} = (s_1, \dots, s_T)$ is a state rollout from passing a control trajectory $\boldsymbol{u} = (u_0, \dots, u_{T-1})$ drawn from a space of controls $\mathcal{U}$ through the dynamics g; $\hat{\boldsymbol{s}}^i = (\hat{s}^i_1, \dots, \hat{s}^i_T)$ is a trajectory prediction for agent i, extracted using f, using past trajectory of horizon h and $\hat{\boldsymbol{s}}^{1:n} = (\hat{\boldsymbol{s}}^1, \dots, \hat{\boldsymbol{s}}^n)$ as input; $\mathcal{J}$ is a cost expressing considerations of safety and efficiency where $\hat{\boldsymbol{s}}^{1:n}$ is an input, so the evaluation of $\boldsymbol{u}$ and thus, the controller performance relies on the quality of f.

B. MPC with Probabilistic Prediction

We integrate the models from Sec. III into a composite loss function where the expectation is taken over the uncertainty over the future human behavior represented by the distribution $(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}^{1:n}) \sim p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}^{1:n} | s_{t-h:t}, s^{1:n}_{t-h:t})$:

$$\begin{aligned} \mathcal{J}^{exp}(\boldsymbol{s}, \hat{\boldsymbol{s}}^{1:n}) = a_g \mathcal{J}_g(\boldsymbol{s}) + \\ \mathbb{E}\left[a_d \mathcal{J}_d(\boldsymbol{s}, \hat{\boldsymbol{s}}^{1:n}) + a_p \mathcal{J}_p(\boldsymbol{s}, \hat{\boldsymbol{s}}^{1:n}) + a_c \mathcal{J}_c(\boldsymbol{s}, \hat{\boldsymbol{s}}^{1:n})\right] \end{aligned}, \tag{2}$$

Prediction inconsistency cost. Minimizing the cost:

$$\mathcal{J}_c(\boldsymbol{s}, \hat{\boldsymbol{s}}) = \mathbb{E}\left[\|\boldsymbol{s} - \hat{\boldsymbol{s}}\|\right], \tag{3}$$

matches in expectation the model prediction for the robot $\hat{\boldsymbol{s}}$, given its past crowd interactions. Aligning with the predictions allows for more confident human predictions and discourages behavior leading to unexpected navigation scenarios.

C. Simulated Experiments

To evaluate the impact of prediction accuracy on navigation performance, we deployed Honda's experimental ballbot [11, 18, 33] (see Fig. 1) in a simulated Gazebo world with three ORCA [30] agents, where they move across the diagonals of a workspace (see Table I, top left). We evaluated navigation performance in terms of *Safety* - minimum distance between the robot and agents throughout a trial, and *Time to goal* - time taken by the robot to reach its goal.

Implementation. We instantiated four different MPC variants, each using a different mechanism for motion prediction: *MPC with CV prediction*, *MPC with CVN prediction*, *MPC with S-GAN-1 prediction* and *MPC with S-GAN-20 prediction*. We follow an implementation similar to Brito et al. [1], extracting a set $\mathcal{U}$ of robot control trajectories by propagating the robot with constant velocity towards subgoals.

Results. Fig. 3a depicts multistep displacement errors across models. We see that S-GAN models' error is consistently higher than CV/CVN. We suspect that this is because ORCA behavior (majorly linear segments) can be approximated using CV-based models. However, we see

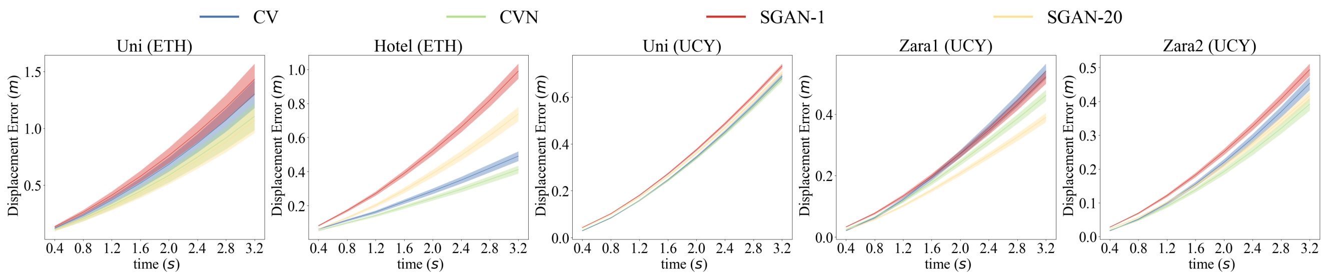
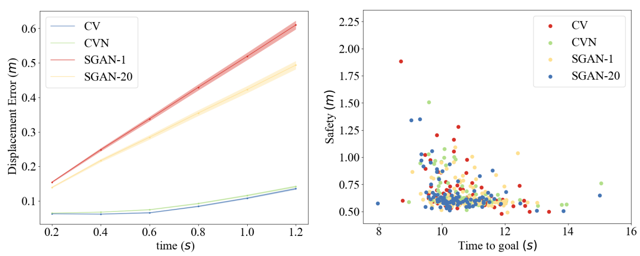


Fig. 2: Error in trajectory prediction of humans on the ETH [21] and UCY [14] datasets. Baselines are referred to from [27]. Error bars indicate 95% confidence intervals, and the line represents the minimum displacement error across the samples.



(a) (b)

Fig. 3: Simulation results. (a) Human motion prediction error over time. (b) Safety vs Time to goal. Lines represent minimum displacement errors across the samples and error bands indicate 95% confidence intervals.

this CV/CVN prediction superiority does not translate to navigation: a scatter plot for safety vs. time to goal (Fig. 3b) does not show a clear winner.

V. Real-World Experiments

As discussed in Sec. I, benchmarking in simulation has limitations. So, we investigate the relationship between prediction and navigation under realistic settings in the lab.

A. Experimental Setup

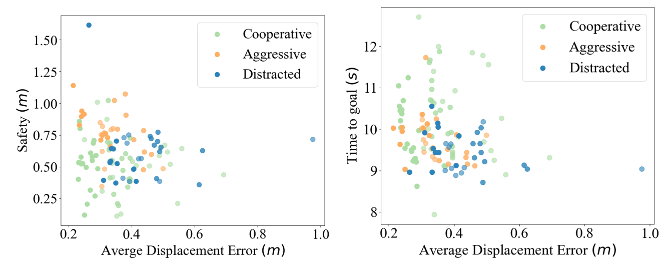
We used Honda's experimental ballbot [11, 18, 33] (see Fig. 1), in a workspace mirroring out simulation setup.

Conditions. We designed three experimental conditions (shown on the left column of Table I) involving robot navigation under different crowd behaviors that a robot could encounter in a crowded space: *Cooperative*, *Aggressive* and *Distracted*. Across all conditions, the robot moves between fixed start and goal points (20 trials per algorithm for the cooperative and 10 trials for the rest).

Algorithms. Across conditions, we compared the performance of the same MPC architecture under three different motion prediction models: CV, S-GAN-1, and S-GAN-20.

Hypotheses. While S-GAN models performed worse than CV in simulation, their prediction accuracy on real-world datasets [8] (Fig. 2) appeared promising for real-world operation. Thus, we expected S-GAN models to outperform baselines and enable improved navigation performance. We formalized these expectations into the following hypotheses:

H1: S-GAN-based prediction is more accurate than CV prediction across all conditions.



(a) *Safety* (b) *Time to goal*

Fig. 4: Relationship between prediction performance and navigation performance per trial in the real world.

H2: S-GAN-based prediction enables the MPC to achieve higher navigation performance across all conditions.

H3: Lower prediction error generally enables the MPC to achieve higher navigation performance.

B. Results

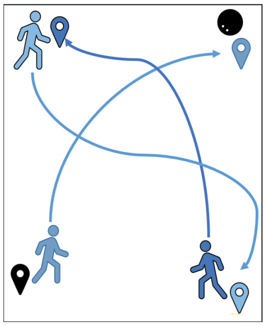
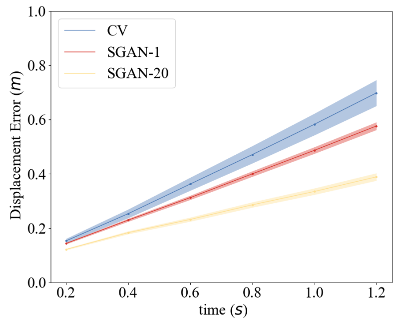
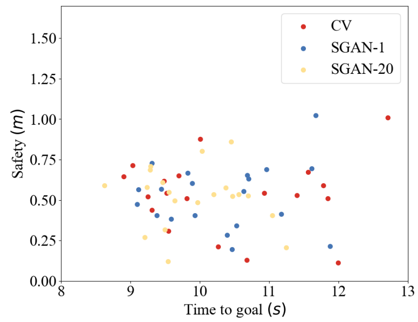
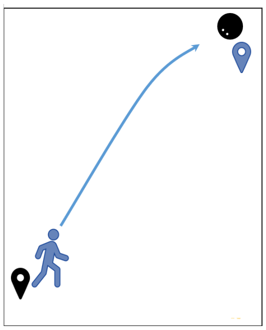
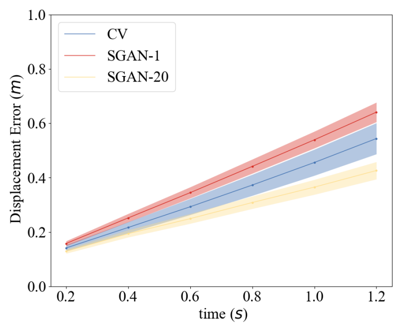
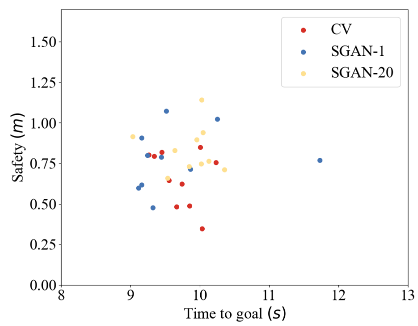
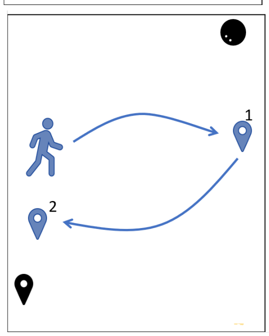
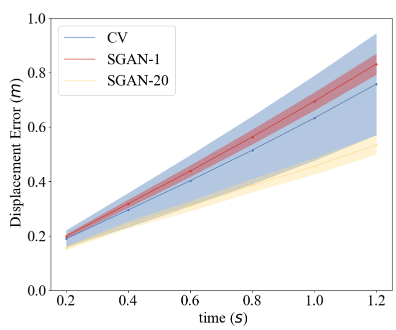
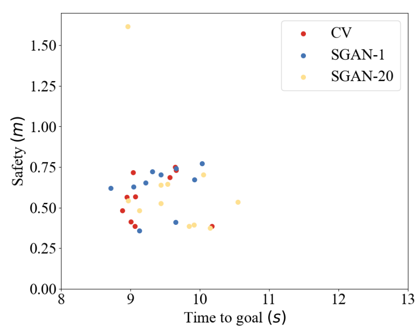
Table I shows the multistep prediction error and the navigation performance distribution per condition. Fig. 4 relates average prediction error per trial to navigation performance.

H1. We see that S-GAN-20 outperforms CV and S-GAN-1 exhibiting consistently lower multistep prediction error (Table I, 2nd column) across all conditions. H1 holds for a strong model like S-GAN-20.

H2. From the right column of Table I, we see that for the cooperative condition, S-GAN-20 has generally is mostly on the left (good time efficiency) and usually higher than the $0.5m$ *Safety*, whereas others are dispersed all over the graph. In the aggressive condition, no major differences in terms of efficiency, but S-GAN-20 is often but not consistently safer than baselines. In the distracted condition, algorithms are close to each other. We find no support that the clear superiority in SGAN-20 predictions (H1) translates to superiority in navigation, and therefore H2 is rejected.

H3. Fig. 4 shows that across conditions, we see a pattern connecting lower errors to higher safety and efficiency. However, this is not definitive: points are scattered across large regions for both metrics. Further, as shown in Table I, prediction rankings do not transfer clearly to navigation rankings. Thus, we find no support that lower prediction error leads to improved navigation and H3 is rejected.

TABLE I: Real-world experiments. Each row shows a different experimental condition: an illustration of the crowd behavior under each condition is shown on the left (users and their goals are shown in blue, whereas the robot and its goal are shown in black color); the multistep prediction error across trials is shown in the middle (error bands indicate 95% confidence intervals); a scatter plot of *Safety* against *Time to goal* is shown on the right.

	Condition	Prediction Error	*Safety* vs *Time to goal*
Cooperative			
Aggressive			
Distracted			

VI. Discussion

Model Transfer. The high-quality predictions of S-GAN transferred from offline datasets to predictions onboard the robot which shows the efficacy of this generative machinery in modeling multiagent interactions. However, S-GANs struggled with out-of-distribution behaviors encountered in the ORCA-simulated trials (Sec. IV-C). This observation highlights the sensitivity of the model to the modes of interaction found in the training dataset. Inducing structure through interaction representations [16, 23, 28] might improve transfer across a wider range of behavior.

Robot and crowd motion are entangled. Across models, prediction performance did not clearly map to navigation performance (H3). During navigation, robot motion is coupled with crowd motion. We accounted for that with a *joint prediction* model, capturing close unfolding crowd-robot interactions. However, when the MPC forces the robot to deviate from the model's ego-prediction, the resulting action likely violates the validity of the crowd motion prediction. While the prediction inconsistency cost (see Sec. IV-B) motivated closeness to ego predictions, the other costs may lead to states outside the model's confidence. An exciting direction for future work is incorporating formalisms of prediction model confidence into decision making.

Beyond the Safety-Efficiency tradeoff. After the lab experiments, users shared that MPC with S-GAN was predictable, safer, and more comfortable, but these values are not reflected in the evaluation metrics. While Safety and Efficiency are extensively used for evaluation in social navigation [19], they miss important attributes of interaction such as human comfort, satisfaction, and smoothness. Recent work has looked at the connection between robot navigation and human impressions [17, 31] but more work is needed on the design of validated interaction metrics, integrating crucial modalities like human gaze, body posture, and gestures.

References

[1] B. Brito, M. Everett, J. P. How, and J. Alonso-Mora. Where to go next: Learning a subgoal recommendation policy for navigation in dynamic environments. *IEEE Robotics and Automation Letters*, 6(3):4616–4623, 2021.

[2] S. Casas, W. Luo, and R. Urtasun. Intentnet: Learning to predict intention from raw sensor data. In *Proceedings of the Conference on Robot Learning (CoRL)*, 2018.

[3] C. Chen, Y. Liu, S. Kreiss, and A. Alahi. Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 6015–6022, 2019.

[4] C. Chen, S. Hu, P. Nikdel, G. Mori, and M. Savva. Relational graph learning for crowd navigation. In *Proceesings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10007–10013, 2020.

[5] N. Djuric, V. Radosavljevic, H. Cui, T. Nguyen, F.-C. Chou, T.-H. Lin, and J. G. Schneider. Motion prediction of traffic actors for autonomous driving using deep convolutional networks. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018.

[6] M. Everett, Y. F. Chen, and J. P. How. Motion planning among dynamic, decision-making agents with deep reinforcement learning. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3052–3059, 2018.

[7] T. Fraichard and V. Levesy. From crowd simulation to robot navigation in crowds. *IEEE Robotics and Automation Letters*, 5(2):729–735, 2020.

[8] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi. Social GAN: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2255–2264, 2018.

[9] D. Helbing and P. Molnár. Social force model for pedestrian dynamics. *Physical Review E*, 51(5):4282–4286, 1995.

[10] S. Hoermann, M. Bach, and K. Dietmayer. Dynamic occupancy grid prediction for urban autonomous driving: A deep learning approach with fully automatic labeling. In *Proceedings of the International Conference on Robotics and Automation (ICRA)*, page 2056–2063, 2018.

[11] Honda. Honda P.A.T.H. Bot, 2019. URL https://global.honda/innovation/CES/2019/path_bot.html.

[12] Y. Kobayashi, T. Sugimoto, K. Tanaka, Y. Shimomura, F. J. A. Garcia, C. H. Kim, H. Yabushita, and T. Toda. Robot navigation based on predicting of human interaction and its reproducible evaluation in a densely crowded environment. *International Journal of Social Robotics*, 14(2):373–387, may 2021.

[13] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard. Socially compliant mobile robot navigation via inverse reinforcement learning. *The International Journal of Robotics Research*, 35(11):1289–1307, 2016.

[14] A. Lerner, Y. Chrysanthou, and D. Lischinski. Crowds by example. *Computer Graphics Forum*, 26(3):655–664, 2007.

[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 *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3517–3524, 2021.

[16] K. Mangalam, Y. An, H. Girase, and J. Malik. From goals, waypoints & paths to long term human trajectory forecasting. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 15213–15222, 2021.

[17] C. Mavrogiannis, P. Alves-Oliveira, W. Thomason, and R. A. Knepper. Social momentum: Design and evaluation of a framework for socially competent robot navigation. *Transactions on Human-Robot Interaction*, 11(2), 2022.

[18] C. Mavrogiannis, K. Balasubramanian, S. Poddar, A. Gandra, and S. S. Srinivasa. Winding through: Crowd navigation via topological invariance. *IEEE Robotics and Automation Letters*, 8(1):121–128, 2023.

[19] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh. Core Challenges of Social Robot Navigation: A Survey. *Transactions on Human-Robot Interaction*, 12(3), 2023.

[20] H. S. Park, J.-J. Hwang, Y. Niu, and J. Shi. Egocentric future localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4697–4705, 2016.

[21] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 261–268, 2009.

[22] S. Poddar, C. Mavrogiannis, and S. S. Srinivasa. From crowd motion prediction to robot navigation in crowds. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[23] J. Roh, C. Mavrogiannis, R. Madan, D. Fox, and S. Srinivasa S. Multimodal trajectory prediction via topological invariance for navigation at uncontrolled intersections. In *Proceedings of the Conference on Robot Learning (CoRL)*, 2020.

[24] 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*, 39(8):895–935, 2020.

[25] T. Salzmann, B. Ivanovic, P. Chakravarty, and M. Pavone. Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 683–700, 2020.

[26] A. J. Sathyamoorthy, J. Liang, U. Patel, T. Guan, R. Chandra, and D. Manocha. DenseCAvoid: Real-

time navigation in dense crowds using anticipatory behaviors. *IEEE International Conference on Robotics and Automation (ICRA)*, pages 11345–11352, 2020.

[27] C. Schöller, V. Aravantinos, F. Lay, and A. Knoll. What the constant velocity model can teach us about pedestrian. *IEEE Robotics and Automation Letters*, 5 (2):1696–1703, 2020.

[28] M. Sun, F. Baldini, P. Trautman, and T. Murphey. Move Beyond Trajectories: Distribution Space Coupling for Crowd Navigation. In *Proceedings of Robotics: Science and Systems (RSS)*, 2021.

[29] P. Trautman, J. Ma, R. M. Murray, and A. Krause. Robot navigation in dense human crowds: Statistical models and experimental studies of human-robot cooperation. *International Journal of Robotics Research*, 34(3):335–356, 2015.

[30] J. van den Berg, S. J. Guy, M. Lin, and D. Manocha. Reciprocal n-body collision avoidance. In *Robotics Research*, pages 3–19. Springer Berlin Heidelberg, 2011.

[31] N. Walker, C. Mavrogiannis, S. S. Srinivasa, and M. Cakmak. Influencing behavioral attributions to robot motion during task execution. In *Proceedings of the Conference on Robot Learning (CoRL)*, 2021.

[32] A. Wang, C. Mavrogiannis, and A. Steinfeld. Group-based motion prediction for navigation in crowded environments. In *Proceedings of the Conference on Robot Learning (CoRL)*, 2021.

[33] K. Yamane and C. Kurosu. Stable balance controller, March 2020. US Patent 16/375,111.

[34] P. Zhang, W. Ouyang, P. Zhang, J. Xue, and N. Zheng. SR-LSTM: State refinement for LSTM towards pedestrian trajectory prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12077–12086, 2019.

[35] B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa. Planning-based prediction for pedestrians. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3931–3936, 2009.