Intent-Aware Pedestrian Prediction for Adaptive Crowd Navigation

Kapil D. Katyal[1,2], Gregory D. Hager[2] and Chien-Ming Huang[2]

***Abstract*— Mobile robots capable of navigating seamlessly and safely in pedestrian rich environments promise to bring robotic assistance closer to our daily lives. In this paper we draw on insights of how humans move in crowded spaces to explore how to recognize pedestrian navigation intent, how to predict pedestrian motion and how a robot may adapt its navigation policy dynamically when facing unexpected human movements. Our approach is to develop algorithms that replicate this behavior. We experimentally demonstrate the effectiveness of our prediction algorithm using real-world pedestrian datasets and achieve comparable or better prediction accuracy compared to several state-of-the-art approaches. Moreover, we show that confidence of pedestrian prediction can be used to adjust the risk of a navigation policy adaptively to afford the most comfortable level as measured by the frequency of personal space violation in comparison with baselines. Furthermore, our adaptive navigation policy is able to reduce the number of collisions by 43% in the presence of novel pedestrian motion not seen during training.**

I. Introduction

As we continue to develop mobile robots to support various human activities ranging from security surveillance to warehouse automation, it is critical for these robots to move efficiently and safely around humans. To successfully integrate these robots into human environments, we built on prior research exploring robot navigation in human crowds (e.g., [1]) and drew insights from seamless human navigation through crowded spaces. One way that people accomplish seamless navigation is by anticipating other pedestrians' future movements and adjusting their own behaviors accordingly; for example, people slow down or change directions to avoid collisions [2]. In this work, we seek to computationally realize and evaluate human-inspired movement anticipation and how such anticipation may enhance the quality of robot navigation in terms of success rate and pedestrian comfort.

Specifically, this work makes the following contributions: (1) a novel approach to pedestrian prediction that combines generative adversarial networks with a probabilistic model of intent that achieves performance which matches or exceeds state-of-the-art baseline algorithms on real world datasets, (2) the ability to use errors in predicted pedestrian motion to detect novel pedestrian behaviors not seen during training, (3) an adaptive policy that adjusts the risk of the robot controller based on detecting novel pedestrian behaviors to minimize collisions.

The remainder of this paper is organized as follows. In Sec. II, we provide a brief overview of existing work on robot navigation in human crowds. In Sec. III, we describe a novel approach to longer term pedestrian prediction that probabilistically reasons about the pedestrian's intent. In Sec. IV, we compare our prediction results using real world datasets and show comparable or better performance than state-of-the-art baselines. In Sec. V, we describe how measuring uncertainty in pedestrian prediction can be used to capture novel pedestrian motion not seen during training. This awareness allows us to adjust the risk of a learned navigation policy resulting in 43% less collisions in the presence of novel pedestrian motion.

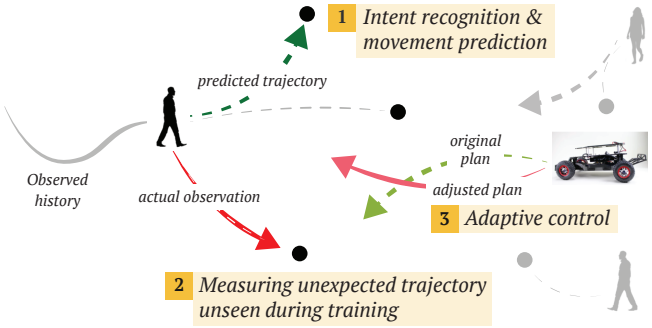


Fig. 1: Adaptive crowd navigation policy that uses pedestrian intent and prediction error to adjust the risk profile of a control policy.

II. Related Work

Crowd Navigation

Previous studies have investigated approaches that enable mobile robot navigation in crowded environments (e.g., [1]). This body of work can be classified into three broad areas: (1) algorithms that react to moving obstacles in real time, (2) trajectory based approaches that plan paths by anticipating future motion of obstacles, and (3) reinforcement learning based approaches that learn a policy to navigate in crowded environments. Reaction based methods include works such as reciprocal velocity obstacles (RVO) [3] and optimal reciprocal collision avoidance (ORCA) [4]. Trajectory based approaches, such as [5], [6], explicitly propagate estimates of future motion over time and perform trajectory optimization on those future states for collision avoidance. Additionally, several recent works use variations of reinforcement learning to learn policies capable of crowd navigation (e.g., [7], [8], [9]). Everett et al. [8] developed a decentralized approach to multiagent collision avoidance using a value network that estimates the time to goal for a given state transition. Chen et al. [9] further extended this work by adding an attention mechanism and a novel pooling method to handle a variable

[1]Johns Hopkins University Applied Physics Lab, Laurel, MD, USA. Kapil.Katyal@jhuapl.edu
[2]Dept. of Comp. Sci., Johns Hopkins University, Baltimore, MD, USA.

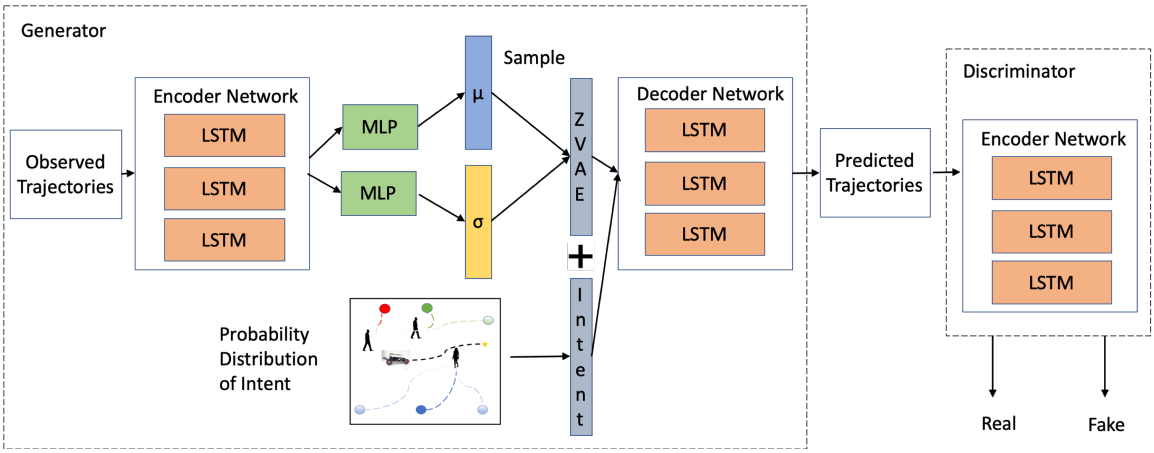


Fig. 2: Our pedestrian prediction network architecture. The generator network consists of a recurrent encoder network, a variational autoencoder, an intent prediction module and recurrent decoder network. The discriminator network consists of a recurrent encoder network to distinguish between real and fake trajectories.

number of humans in the scene. Kahn et al. [10] investigates adaptive navigation polices based on uncertainty; however, they only considered environment uncertainty with static obstacles and not navigation in the presence of pedestrians.

Pedestrian Prediction

Several studies have investigated pedestrian prediction for a variety of applications including robotics, autonomous driving, and video surveillance. Many approaches treat pedestrian prediction as a state estimation problem by relying on a kinematic model and using concepts from Bayesian and Kalman Filtering [11], [12], [13], [14]. Many other works have investigated intent or goal based estimation as part of trajectory planning (e.g., [15], [16], [17]). Recent works have investigated deep neural networks that consider agent-to-agent and agent-to-environment interactions (e.g., [18], [19], [20], [2]). This brief summary only highlights a small snapshot of the many relevant works related to pedestrian motion prediction. For a more comprehensive overview, Rudenko et al. provide a survey describing various approaches to the human motion trajectory prediction problem [21].

While each of these papers makes significant contributions to their respective fields, none of the prior work, as far as we are aware, focuses on pedestrian prediction as a measure of uncertainty to inform adaptive crowd navigation policies.

III. Pedestrian Prediction

Our objective is to estimate future pedestrian trajectory given an observation history of past trajectories. Formally, at given time t, we represent the state of the pedestrian, i as: $X_i^t = (x_i^t, y_i^t)$. We observe pedestrian state during a time window $t = 1$ to $t = T_{obs}$ represented as $X_i^{obs} = [(x_i^1, y_i^1), ..., (x_i^{T_{obs}}, y_i^{T_{obs}})]$. The objective is to predict the pedestrian state from time $t = T_{obs+1}$ to $t = T_{pred}$ represented as $Y_i^{pred} = [(x_i^{T_{obs+1}}, y_i^{T_{obs+1}}), ..., (x_i^{T_{pred}}, y_i^{T_{pred}})]$.

Neural Network Architecture

Our network architecture described in Fig. 2 consists of a generator and a discriminator network. The generator network includes a recurrent encoder network, a variational autoencoder, a recurrent decoder network, and an intent predictor. The discriminator network classifies samples as either real trajectories or not socially acceptable[18].

Encoder Network

The recurrent encoder network consists of a linear spatial embedding layer with a ReLU activation layer, $\alpha(\cdot)$, followed by an LSTM layer, where W_e and W_{lstme} are weights of the embedding layer and LSTM, respectively.

$$h_i^t = \alpha(X_i^t, W_e)$$
$$p_i^t = LSTM_e(p_i^{t-1}, h_i^t, W_{lstme})$$

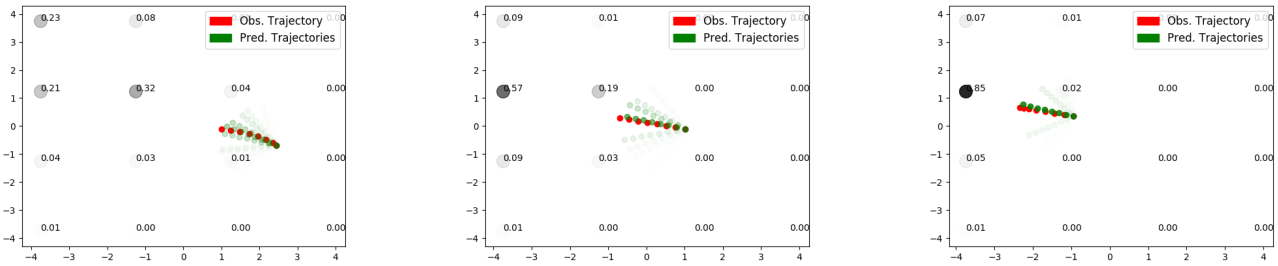
We then use two fully connected linear layers (MLPs) to produce latent distribution parameters μ and σ, and sample from this distribution represented as z_{vae} to generate sample diversity, where W_μ, W_σ, b_μ, and b_σ are the weights and biases of the fully connected layers, respectively.

$$\mu_i^t = W_\mu p_i^t + b_\mu$$
$$\sigma_i^t = \log(\exp(W_\sigma p_i^t + b_\sigma) + 1)$$
$$z\_vae_i^t = \mu_i^t + \sigma_i^t \cdot \epsilon; \epsilon \sim \mathcal{N}(0, 1)$$

Intent Recognition

In this section, we describe our probabilistical model of intent recognition and how navigation intent can be combined with latent embeddings from the recurrent encoder network to improve longer term prediction of pedestrian trajectory.

Inspired by [22], we use a Bayesian approach to estimate the probability of a desired goal, g_i^t, of the pedestrian i based on a past observation history, X_i^{obs} spanning $t = 1$ to $t = T_{obs}$, as described by Equation 1. $P(g_i^t|X_i^{obs})$ is the posterior probability of each goal, g given an observation history X_i^{obs}. $P(g_i^t)$ is the prior probability of a pedestrian i choosing a given goal at time t. $P(X_i^{obs}|g_t^t)$ is the likelihood probability of observing X_i^{obs} given g_t^t and modelled as a



(a) Start of Trajectory (b) Middle of Trajectory (c) End of Trajectory

Fig. 3: Estimating probability distribution of the pedestrian's intent from the start, middle and end of the trajectories. Here we demonstrate the intent uncertainty decreases as the pedestrian reaches the target.

Gibbs measure described by Equation 2. Here, $E(X_i^{obs}|g_i^t)$ is an energy function that we set equal to distance between the observed trajectory and the shortest trajectory to the goal. β is a hyperparameter that adjusts the landscape of the resulting probability distribution and $Z(\beta)$ is a normalizing constant.

$$P(g_i^t|X_i^{obs}) = \frac{P(g_i^t)P(X_i^{obs}|g_i^t)}{P(X_i^{obs})} \propto P(g_i^t)P(X_i^{obs}|g_i^t) \quad (1)$$

$$P(X_i^{obs}|g_i^t) = \frac{1}{Z(\beta)}\exp(-\beta E(X_i^{obs}|g_i^t)) \quad (2)$$

In our work, we assume $P(g_i^t)$ is a uniform distribution and we set β to 0.5. We explored two ways to represent a discrete set of goals: (1) inferring the goals directly from the data and (2) generating a uniform set of goals in a grid-like pattern. We decided on using a grid-like representation of goals as it affords greater generalization to diverse scenarios. In this work, we generated a uniform 4×4 grid of targets to illustrate our approach. Figs. 3a, 3b, and 3c show the distribution landscape across the goals based on the observed trajectory in red. Green trajectories represent ideal paths to each of the targets with the opacity equal to the resulting probability $P(g_i^t|X_i^{obs})$. As shown in Fig. 3, there initially is significant uncertainty regarding the end goal of the pedestrian; however, as the pedestrian navigates closer to the target, the uncertainty decreases towards a specific goal.

The resulting 16 dimensional probability distribution is concatenated to the z_i^t feature vector sampled from the VAE prior to the decoder network of the generator.

$$z_i^t = [z\_vae_i^t, P(g_i^t|X_i^{obs})]$$

Decoder Network

The goal of the decoder network is to use the latent embedding from the encoder network combined with the intent distribution to generate prediction samples. The decoder network used is similar to that in [18] and consists of linear and recurrent layers used to generate pedestrian predictions. The hidden state of the LSTM decoder, p_i^t is initialized to z_i^t. Here W_d, W_{lstmd} and W_o are weights, and ϕ is a fully connected layer.

$$dh_i^t = \alpha(X_i^{t-1}, W_d)$$

$$p_i^t = LSTM_d(p_i^{t-1}, dh_i^t, W_{lstmd})$$

$$\hat{X}_i^t = \phi(p_i^t, W_o)$$

Loss Functions

During training, we use a combination of the mean squared error (MSE) between the ground truth trajectory and the predicted trajectory, adversarial loss, as well as KL-divergence loss from a unit Gaussian distribution for the variational autoencoder as our loss function. Similar to [18], we also experiment with variety loss during training where for each scene we generate k possible outputs and choose the lowest MSE to generate diversity in the samples.

IV. Pedestrian Prediction Experiments

We performed a series of experiments with real world datasets and compared our performance with several state-of-the-art baseline algorithms. In particular, we used two widely used, publicly available repositories—ETH [23] and UCY [24]—consisting of 5 unique datasets (ETH, Hotel, Univ, Zara1, and Zara2) with 4 scenes. The datasets include pedestrian motion with a top down view and annotated pedestrian positions with respect to the world frame.

Training Details

The dimension of the hidden state for the encoder, p is 16. The dimension of the decoder network's hidden state, z is 32 including the 16 dimensions representing probability of navigation intent. We trained for 200 epochs with a batch size of 64 using the Adam optimizer with the initial generator network learning rate set to 0.0001 and the initial discriminator network learning rate set to 0.001.

Baseline Implementations

We compare our algorithm to several baselines representing unique solutions to the pedestrian prediction problem. These baselines include a linear regression algorithm (Linear) that minimizes least square error to estimate parameters of a linear model, a simple LSTM model, (S-LSTM) which

is an LSTM model with a social pooling layer [25], Social GAN (SGAN) which uses a GAN in combination with an S-LSTM [18], and finally SoPhie which uses scene contextual information to make predictions [19].

Metrics

Similar to [18], [19], [26], we use the average displacement error (ADE) and final displacement error (FDE) metrics to compare our approach to existing baselines. The ADE (in meters) is the L2 distance between the ground truth and predicted pedestrian trajectories for each trajectory point. The FDE (in meters) is the final displacement distance between the last point in the predicted trajectory and the ground truth. We use a leave-one-out approach where we train on 4 of the datasets and test on one.

Prediction Results

Table I summarizes the ADE and the FDE for the 5 datasets described above when observing 8 timesteps (3.2 sec) and predicting 12 timesteps (4.8 sec) into the future. We use similar notation from [18], kV(+IR)-N, where k represents the samples for variety loss, N represents the number of samples taken during test time and +IR represents whether intent recognition is used as part of the prediction. Sample diversity in this sense [18] allows the learning algorithm to produce k predictions and choose the prediction with the smallest MSE to encourage diversity. We compare our results with and without sample diversity (where $k = N = 1$). Our approach results in better than or equal ADE and FDE when compared to the baselines for 8 out of 10 experiments without sample diversity and 7 out of the 10 experiments with sample diversity. In addition, our approach achieves best performance when averaging across the 5 datasets for both the ADE and FDE. These results demonstrate the advantages of estimating a probabilistic interpretation of intent and explicitly using this estimate when making longer term predictions of pedestrian trajectories.

V. Adaptive Crowd Navigation

We now extend our pedestrian prediction algorithm to enable adaptive crowd navigation policies for mobile robots. We conjecture that errors in pedestrian prediction can serve as a measure of policy uncertainty. Specifically, as we encounter distribution shift of pedestrian motion, the error in pedestrian prediction can serve as an effective method to detect novel pedestrian motions not seen during training. We believe that detecting novel pedestrian motion profiles is a cue to switch to a risk averse control policy and by doing so will result in fewer collisions.

In our development of adaptive navigation, we leverage the CrowdSim simulation environment provided by [9]. This environment provides the ability to model pedestrian motion using an optimal reciprocal collision avoidance (ORCA) model [4]. Pedestrian behavior can be modelled using parameters such as preferred velocity, the maximum distance and time to take into account neighboring agent behavior, pedestrian's radius, and maximum velocity. In addition, CrowdSim provides an OpenAI Gym like environment [27] to experiment with reinforcement learning based policies controlling a robot's actions to reach a target goal while avoiding obstacles.

The state space of the environment follows that of [28], [9] and consists of the following parameters with respect to the robot position as the origin and the x-axis pointing towards the goal: distance from robot position to goal, robot's preferred velocity, actual velocity and radius. For each pedestrian, the state includes position, velocity, radius, and distance between pedestrian and robot.

The action space assumes a nonholonomic unicycle kinematic model for the robot agent and consists of 3 discrete speeds and 6 discrete rotation angles for a total of 18 actions.

In order to learn a policy that allows the robot to successfully reach the target while avoiding collisions with other pedestrians, we use the same reward definition as [8], [9]:

$$R(\mathbf{s},\mathbf{a}) = \begin{cases} -0.25 & \text{if} \quad d_{min} < 0 \\ -0.1 - d_{min}/2 & \text{else if} \quad d_{min} < 0.2 \\ 1 & \text{else if} \quad \text{robot reached goal} \\ 0 & \text{o.w.} \end{cases}$$

where d_{min} is the minimum distance separating the robot and the humans during the previous timestep.

We train two navigation policies, a risk averse and an aggressive policy. The aggressive policy consists of a preferred velocity of 2.0 m/s, and the risk averse policy is limited to 1.0 m/s [1]. We first assess state-of-the-art crowd navigation policies' performance in the presence of a changing distribution of pedestrian motion. We trained both CADRL [28] and SARL [9] with preferred robot velocities set to 2.0 m/s with a static pedestrian motion model. The starting and ending positions of the pedestrians were sampled uniformly inside a square of width 10 meters. The policies were pretrained using imitation learning using ORCA similar to [9] and were subsequently trained using a value iteration network for 10000 steps. We then evaluated the performance of these policies on test data with a shifted distribution of pedestrian motion uniformly sampled from the parameters described in Table II. The objective for this test is to show the limitation in prior works' ability to handle distribution shifts from unseen pedestrian motion during training.

We then conducted a series of experiments with an adaptive control policy using various methods of novelty detection of new pedestrian behaviors.

We first evaluate with a traditional, non-deep learning based approach using a one-class SVM with a radial basis kernel [29]. We train the one-class SVM algorithm based on the fixed pedestrian motion profile and evaluate its ability to detect novel distributions of pedestrian motion data. We then conduct several experiments using deep learning based approaches for novelty detection including Social GAN and

[1]For the risk averse policy, we first considered training a policy where the penalty for collision was significantly increased, however, changing the reward function would make comparisons with existing work difficult.

Metric	Dataset	Linear	LSTM	S-LSTM	Sophie	SGAN				Ours			
					1V-20	1V-1	1V-20	20V-20	20VP-20	1V-1	1V+IR-1	1V+IR-20	20V+IR-20
ADE	**ETH**	1.33	1.09	1.09	0.70	1.13	1.03	0.81	0.87	0.96	0.85	0.77	**0.69**
	HOTEL	**0.39**	0.86	0.79	0.76	1.01	0.90	0.72	0.67	0.60	0.48	0.42	**0.39**
	UNIV	0.82	0.61	0.67	0.54	0.60	0.58	0.60	0.76	0.55	0.53	**0.51**	0.56
	ZARA1	0.62	0.41	0.47	**0.30**	0.42	0.38	0.34	0.35	0.45	0.41	0.36	0.35
	ZARA2	0.77	0.52	0.56	0.38	0.52	0.47	0.42	0.42	0.38	0.33	**0.30**	0.31
	AVG	0.79	0.70	0.72	0.54	0.74	0.67	0.58	0.61	0.59	0.52	0.47	**0.46**
FDE	**ETH**	2.94	2.41	2.35	1.43	2.21	2.02	1.52	1.62	1.85	1.80	1.66	**1.42**
	HOTEL	**0.72**	1.91	1.76	1.67	2.18	1.97	1.61	1.37	1.18	1.04	0.94	0.79
	UNIV	1.59	1.31	1.40	1.24	1.28	1.22	1.26	1.52	1.17	1.13	**1.09**	1.17
	ZARA1	1.21	0.88	1.00	**0.63**	0.91	0.84	0.69	0.68	0.94	0.87	0.79	0.74
	ZARA2	1.48	1.11	1.17	0.78	1.11	1.01	0.84	0.84	0.79	0.72	**0.65**	0.66
	AVG	1.59	1.52	1.54	1.15	1.54	1.41	1.18	1.21	1.19	1.11	1.03	**0.96**

TABLE I: Average Displacement Error (ADE) and Final Displacement Error (FDE) in meters for $t_{pred} = 12$ timesteps. Our method matches or outperforms state-of-the-art and baseline methods by explicitly estimating intent as a probability distribution of possible goals (lower is better).

our intent-aware pedestrian prediction algorithm. Our goal is to demonstrate the benefits that a higher performing pedestrian prediction algorithm can have on reducing collisions in an adaptive crowd navigation policy.

We trained both Social GAN and our prediction algorithm with the same fixed pedestrian motion profile that we trained the original policies. We then tested pedestrian prediction with a changing distribution of pedestrian motion while allowing the robot to navigate. We compute an estimate of novelty by thresholding the prediction error, as measured using the FDE, by a value α. If the error exceeds α, the policy moves from an aggressive behavior to a risk averse policy with the goal of avoiding collisions. The value of α was chosen by computing the mean and standard deviation of the FDE in the training set. In our experiments, α was set to 3 standard deviations from the mean to eliminate outliers.

TABLE II: ORCA Model Parameters

Parameters Name	Min Value	Max Value
Preferred Velocity	0.5	2.0
Radius	0.2	0.8
Neighbor Distance	2.0	20.0
Time Horizon	0.1	5.0

A. Quantitative Analysis

The primary metrics we used for comparison are the number of successful trials, number of collisions, the average navigation time, the discomfort level, and the average reward. The discomfort level is defined as the frequency of the separating distance being less than the desired separation distance, in this case 0.2 m. The results after running 500 test cases are reported in Table III. We compare various methods of crowd navigation denoted by $method - p$ where p is the number of pedestrians in the scene and A indicates an adaptive policy. In these experiments, the starting and goal positions, and pedestrian motion profiles are all randomly initialized for each of the 500 test cases; however, these parameters are consistent across the various methods to create a fair comparison.

The first two rows show the performance of CADRL and SARL policies where we train and test without changing the distribution of the pedestrian motion profile. We then evaluate the algorithms' ability to avoid collisions when faced with novel pedestrian motion and show that both the CADRL and SARL policies have a significantly higher collision and discomfort rates. The number of collisions for CADRL and SARL increase by 35 and 65, respectively.

The non-deep learning based approach detects novel pedestrian motion using a one-class SVM and reduces the number of collisions by 2 compared to the non-adaptive SARL algorithm.

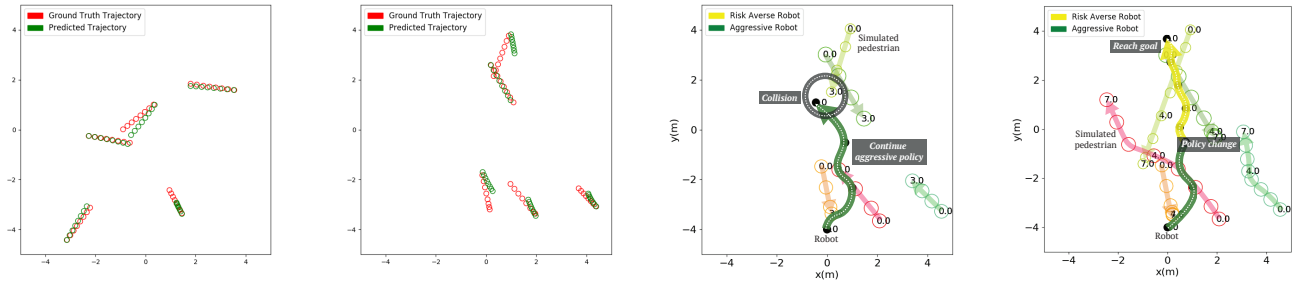
Using Social GAN as the pedestrian prediction algorithm for novelty detection had a significant improvement compared to the one-class SVM by further reducing the number of collisions by 20.

Our intent-aware pedestrian prediction algorithm for novelty detection resulted in best performance across almost all of the metrics. Using our approach, we were able to further reduce the number of collisions by 5 compared to Social GAN and overall by 30 compared to the non-adaptive SARL policy. We do this while also demonstrating best performance in overall discomfort rate and overall reward. Further, we show that the benefits of our approach scale as the number of pedestrians in the scene increase.

B. Qualitative Analysis

We further study the qualitative aspects of our approach. In Fig 4a, we show an example where our pedestrian prediction algorithm has low FDE hence has high confidence and correctly decides to maintain a high risk, fast navigation policy. Conversely, in Fig. 4b, we show an example where the FDE in pedestrian motion is high. This situation is flagged as a novel pedestrian behavior and in this example, a risk averse policy is selected resulting in an avoided collision.

In Fig 4c and Fig. 4d, we also show example trajectories of the adaptive and non-adaptive policies. In Fig 4c, with the non-adaptive policy, even though the pedestrian prediction error is high, the default aggressive policy continues and the robot eventually collides with a pedestrian within 3



(a) Example prediction with low uncertainty

(b) Example prediction with high uncertainty

(c) Example of non-adaptive policy colliding with pedestrians

(d) Example of adaptive policy successfully reaching target while avoiding pedestrians

Fig. 4: This figure provides representative examples of pedestrian predictions with low and high uncertainty as well as the trajectories of the robot and pedestrians navigating to their desired goals.

seconds. In contrast, using our adaptive policy as shown in Fig. 4d, we detect novel pedestrian behavior and instead modify the policy to a risk averse controller. This adaptation causes the robot to reduce its velocity in real time preventing a near collision from occurring. Videos of this behavior and additional examples can be found in the supplemental material of this paper.

TABLE III: Quantitative Analysis of Collision Avoidance

Method	Dist. Shift	Succ.	Coll.	Time Out	Nav. Time	Disc. Rate	Avg. Rwd.
CADRL-5	N	455	45	0	4.48	2.02	0.349
SARL-5	N	490	5	5	4.61	0.99	0.389
CADRL-5	Y	420	80	**0**	**4.52**	3.53	0.296
SARL-5	Y	425	70	5	4.62	2.27	0.303
SVM-A-5	Y	426	68	6	5.34	2.14	0.331
SGAN-A-5	Y	445	45	10	6.31	2.03	0.386
Ours-A-5	Y	**450**	**40**	10	6.74	**1.98**	**0.409**
SARL-10	Y	388	99	13	**5.21**	2.62	0.234
Ours-A-10	Y	**444**	**54**	**2**	8.49	**2.18**	**0.330**
SARL-15	Y	290	205	5	**5.30**	4.89	0.115
Ours-A-15	Y	**366**	**132**	**2**	8.69	**4.20**	**0.212**
SARL-20	Y	172	324	4	**5.27**	6.70	-0.017
Ours-A-20	Y	**262**	**237**	**1**	8.65	**6.29**	**0.066**

VI. Discussion

To enhance robot navigation in human crowds, in this paper we propose techniques that can estimate pedestrian motion and use this motion to enable adaptive policies based on detecting pedestrian movement profiles not present during training. When unexpected pedestrian motions are observed, we modify the policy to a low risk controller with the goal of avoiding collisions. In particular, we present empirical evidence (Table I) showing the importance of having an explicit, probabilistic representation of the intent in longer term prediction. Our pedestrian prediction method performs comparably, if not better, when compared to several state-of-the-art baselines using standard real world datasets. Moreover, while showing how several crowd navigation methods fail to avoid collisions in the presence of novel pedestrian motion not seen during training, our approach demonstrated the best results in terms of most number of successful trials, highest overall reward, lowest number of collisions, and lowest overall discomfort rate.

Our approach is not without limitations. First, our average navigation time did increase compared to the other approaches, although this increase was expected as the reported times excluded test cases where a collision occurred. Moreover, our approach avoided collisions by reducing the speed of the robot which resulted in an expected increase in average navigation time. Second, while our approach reduced the number of collisions, it did not prevent them entirely. There were instances when even though our prediction was accurate, a collision could still occur, suggesting that pedestrian prediction alone does not capture all aspects of uncertainty involved in dynamic crowd navigation. In these scenarios, it may be beneficial to consider other forms of uncertainty estimation such as bootstrapping [30], stochastic dropout [31], [32], and multiple hypothesis loss techniques [33], [34]. Third, this work only focused on modeling pedestrians' navigation intent to enhance the quality of robot navigation. Future work should consider other aspects of pedestrian navigation, such as personality [35] and group interaction, to capture nuances in human navigation. Finally, while detecting novel states is a step forward, our goal, ultimately is to develop continual learning approaches that learn in real time from experiences. We leave this for future work as well.

In spite of these limitations, we believe explicitly modelling a probabilistic interpretation of intent has shown to improve accuracy of estimating future pedestrian motion. Further, the use of pedestrian prediction has the ability to detect novel pedestrian behaviors not seen while learning a policy. Finally, we show that detecting novel pedestrian behaviors and adapting the policy can significantly reduce the number of collisions compared to alternative approaches.

Acknowledgments

This work was partially funded by NSF grant no. 1637949.

REFERENCES

[1] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1726–1743, 2013.

[2] A. Vemula, K. Mülling, and J. Oh, "Modeling cooperative navigation in dense human crowds," *CoRR*, vol. abs/1705.06201, 2017. [Online]. Available: http://arxiv.org/abs/1705.06201

[3] J. P. van den Berg, M. C. Lin, and D. Manocha, "Reciprocal velocity obstacles for real-time multi-agent navigation," *2008 IEEE International Conference on Robotics and Automation*, pp. 1928–1935, 2008.

[4] J. van den Berg, S. Guy, M. Lin, and D. Manocha, *Reciprocal n-Body Collision Avoidance*, 04 2011, vol. 70, pp. 3–19.

[5] M. Phillips and M. Likhachev, "Sipp: Safe interval path planning for dynamic environments," 06 2011, pp. 5628 – 5635.

[6] 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, Jul 2013. [Online]. Available: https://doi.org/10.1007/s10514-013-9334-3

[7] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, "Socially compliant mobile robot navigation via inverse reinforcement learning," *The International Journal of Robotics Research*, vol. 35, no. 11, pp. 1289–1307, 2016. [Online]. Available: https://doi.org/10.1177/0278364915619772

[8] M. Everett, Y. F. Chen, and J. P. How, "Motion planning among dynamic, decision-making agents with deep reinforcement learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, Sep. 2018. [Online]. Available: https://arxiv.org/pdf/1805.01956.pdf

[9] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," 2018.

[10] G. Kahn, A. Villaflor, V. Pong, P. Abbeel, and S. Levine, "Uncertainty-aware reinforcement learning for collision avoidance," *ArXiv*, vol. abs/1702.01182, 2017.

[11] V. Karasev, A. Ayvaci, B. Heisele, and S. Soatto, "Intent-aware long-term prediction of pedestrian motion," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 2543–2549.

[12] A. Elnagar, "Prediction of moving objects in dynamic environments using kalman filters," in *Proceedings 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation (Cat. No.01EX515)*, July 2001, pp. 414–419.

[13] A. Barth and U. Franke, "Where will the oncoming vehicle be the next second?" in *2008 IEEE Intelligent Vehicles Symposium*, June 2008, pp. 1068–1073.

[14] T. Batz, K. Watson, and J. Beyerer, "Recognition of dangerous situations within a cooperative group of vehicles," in *2009 IEEE Intelligent Vehicles Symposium*, June 2009, pp. 907–912.

[15] Y. Kuwata, J. Teo, G. Fiore, S. Karaman, E. Frazzoli, and J. P. How, "Real-time motion planning with applications to autonomous urban driving," *IEEE Transactions on Control Systems Technology*, vol. 17, no. 5, pp. 1105–1118, Sep. 2009.

[16] M. Luber, J. A. Stork, G. D. Tipaldi, and K. O. Arras, "People tracking with human motion predictions from social forces," in *2010 IEEE International Conference on Robotics and Automation*, May 2010, pp. 464–469.

[17] H. Xue, D. Q. Huynh, and M. Reynolds, "Bi-prediction: Pedestrian trajectory prediction based on bidirectional lstm classification," in *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Nov 2017, pp. 1–8.

[18] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: socially acceptable trajectories with generative adversarial networks," *CoRR*, vol. abs/1803.10892, 2018. [Online]. Available: http://arxiv.org/abs/1803.10892

[19] A. Sadeghian, V. Kosaraju, A. Sadeghian, N. Hirose, and S. Savarese, "Sophie: An attentive GAN for predicting paths compliant to social and physical constraints," *CoRR*, vol. abs/1806.01482, 2018. [Online]. Available: http://arxiv.org/abs/1806.01482

[20] F. Bartoli, G. Lisanti, L. Ballan, and A. D. Bimbo, "Context-aware trajectory prediction," *2018 24th International Conference on Pattern Recognition (ICPR)*, pp. 1941–1946, 2017.

[21] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *ArXiv*, vol. abs/1905.06113, 2019.

[22] S. Qi and S. Zhu, "Intent-aware multi-agent reinforcement learning," *CoRR*, vol. abs/1803.02018, 2018. [Online]. Available: http://arxiv.org/abs/1803.02018

[23] S. Pellegrini, A. Ess, and L. Van Gool, "Improving data association by joint modeling of pedestrian trajectories and groupings," in *Computer Vision – ECCV 2010*, K. Daniilidis, P. Maragos, and N. Paragios, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 452–465.

[24] A. Lerner, Y. Chrysanthou, and D. Lischinski, "Crowds by example," *Computer Graphics Forum*, vol. 26, no. 3, pp. 655–664, 2007. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2007.01089.x

[25] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social lstm: Human trajectory prediction in crowded spaces," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[26] N. Lee, W. Choi, P. Vernaza, C. B. Choy, P. H. S. Torr, and M. K. Chandraker, "DESIRE: distant future prediction in dynamic scenes with interacting agents," *CoRR*, vol. abs/1704.04394, 2017. [Online]. Available: http://arxiv.org/abs/1704.04394

[27] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," *CoRR*, vol. abs/1606.01540, 2016. [Online]. Available: http://arxiv.org/abs/1606.01540

[28] Y. F. Chen, M. Liu, M. Everett, and J. How, "Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning," 05 2017, pp. 285–292.

[29] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *J. Mach. Learn. Res.*, vol. 9, pp. 1871–1874, Jun. 2008. [Online]. Available: http://dl.acm.org/citation.cfm?id=1390681.1442794

[30] B. Efron, *The Jackknife, the bootstrap and other resampling plans*, ser. CBMS-NSF Reg. Conf. Ser. Appl. Math. Philadelphia, PA: SIAM, 1982, lectures given at Bowling Green State Univ., June 1980. [Online]. Available: https://cds.cern.ch/record/98913

[31] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *Proceedings of The 33rd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. F. Balcan and K. Q. Weinberger, Eds., vol. 48. New York, New York, USA: PMLR, 20–22 Jun 2016, pp. 1050–1059. [Online]. Available: http://proceedings.mlr.press/v48/gal16.html

[32] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014. [Online]. Available: http://dl.acm.org/citation.cfm?id=2627435.2670313

[33] C. Rupprecht, I. Laina, R. Dipietro, M. Baust, F. Tombari, N. Navab, and G. Hager, "Learning in an uncertain world: Representing ambiguity through multiple hypotheses," in *Proceedings - 2017 IEEE International Conference on Computer Vision, ICCV 2017*, vol. 2017-October. Institute of Electrical and Electronics Engineers Inc., 12 2017, pp. 3611–3620.

[34] K. Katyal, K. Popek, C. Paxton, P. Burlina, and G. D. Hager, "Uncertainty-aware occupancy map prediction using generative networks for robot navigation," in *2019 International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 5453–5459.

[35] A. Bera, T. Randhavane, and D. Manocha, "Aggressive, tense or shy? identifying personality traits from crowd videos." in *IJCAI*, 2017, pp. 112–118.