



HAL
open science

Machine Learning for HRI: Bridging the Gap between Action and Perception

Oliver Roesler, Elahe Bagheri, Amir Aly

► **To cite this version:**

Oliver Roesler, Elahe Bagheri, Amir Aly (Dir.). Machine Learning for HRI: Bridging the Gap between Action and Perception. 2022. hal-03758051

HAL Id: hal-03758051

<https://hal.science/hal-03758051v1>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Proceedings of the Workshop

Machine Learning for HRI:
Bridging the Gap between
Action and Perception

in Conjunction with the 31st IEEE
International Conference on Robot &
Human Interactive Communication
(Ro-Man 2022)

Oliver Roesler, Elahe Bagheri, and
Amir Aly

User Command Correction for Safe Remote Manipulation in Dynamic Environments

Mincheul Kang¹, Minsung Yoon¹ and Sung-Eui Yoon²

Abstract—Collision avoidance is an important factor for safe robot movement. In remote manipulation, a user’s role is huge in avoiding obstacles because a robot follows a user’s command. Especially, dynamic obstacle avoidance requires more user judgment. Unfortunately, a human sometimes decides an unrealizable command with the possibility of collision with obstacles. To cope with the unrealizable command, we present a learning-based user command correction method. Our method predicts the risk of dynamic obstacles and corrects a user’s unrealizable command to avoid collision risk with dynamic obstacles. In this paper, we define the problem to be solved and introduce the proposed method briefly.

I. INTRODUCTION

Remote manipulation has been used to perform dangerous and sophisticated tasks, such as works in nuclear power plants [1] and telesurgery [2], instead of humans. Recently, it has been expanding the scope to environments around us. For example, Telexistence¹ has been researching for remote manipulation to reduce the workforce going to convenience stores far away. In environments around us like a convenience store, there are various and dynamic obstacles. Therefore, handling such obstacles in the progress of manipulation planning is important.

Motion planning essentially considers avoiding collision with obstacles because it is an important factor in terms of safety. Although most motion planners can compute a collision-free trajectory for static obstacles, there is no way to guarantee collision avoidance for dynamic obstacles [3]. Therefore, dynamic obstacle avoidance is a difficult problem and important for safe robot movement.

Recent inverse kinematics (IK) methods have been proposed for remote manipulation with handling dynamic obstacles. CollisionIK [4] quickly computes the distance between robot links and dynamic obstacles by applying a fast convex shape representation method. The computed distances are applied to their optimizer to solve the IK problem for remote manipulation. RCIK [5] is a sampling-based approach that generates IK candidates and then selects one IK candidate away from obstacles. This method utilizes deep learning to quickly compute collision costs of IK candidates from sensor data in real-time.

Even though these methods consider avoiding dynamic obstacles, they depend more on a user’s judgment as it has

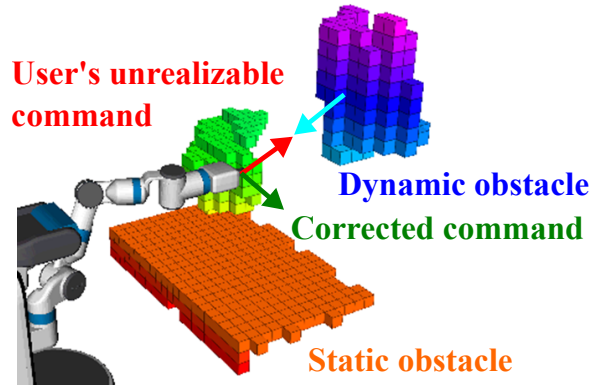


Fig. 1. Our goal is to correct a user’s unrealizable command with a collision risk of dynamic obstacles to moving to a safe region.

to follow a user’s command. In particular, sampling-based IK methods like RCIK are more affected because they aim to follow a user’s commands exactly. In the case of static obstacles, even if a user’s command is wrong, an IK solver considering collision avoidance can stop a robot to ensure safety. On the other hand, for dynamic obstacles, simply stopping is not enough. Since a human does not always decide a perfect command, we need a way to correct a user’s unrealizable command that has the risk of collision with dynamic obstacles.

In this paper, we present a learning-based user command correction method in dynamic environments for safe remote manipulation (Fig. 1). Our method consists of two networks, a risk prediction network (RPN) for dynamic obstacles and a command correction network (CCN). The RPN predicts the risk of dynamic obstacles from successive occupancy grids and joint positions, and the CCN generates a command to move away from the obstacle. According to the predicted risk, we determine whether to follow the user’s command or the CCN’s command.

II. APPROACH

In this section, we introduce a problem definition and brief our approach.

A. Problem definition

In our problem, we deal with remote manipulation where a user consecutively gives a command, $x_u \in \mathbb{R}^6$, to the end-effector in Cartesian space. For the command, a real-time IK solver, such as RCIK, synthesizes a joint configuration, q . We also target a redundant manipulator with multiple joint

¹Mincheul Kang (mincheul.kang@kaist.ac.kr),
¹Minsung Yoon (minsung.yoon@kaist.ac.kr) are with the School of Computing, and ²Sung-Eui Yoon (Corresponding author, sungeui@kaist.edu) is with the Faculty of School of Computing, KAIST at Daejeon, Korea 34141

¹<https://tx-inc.com>

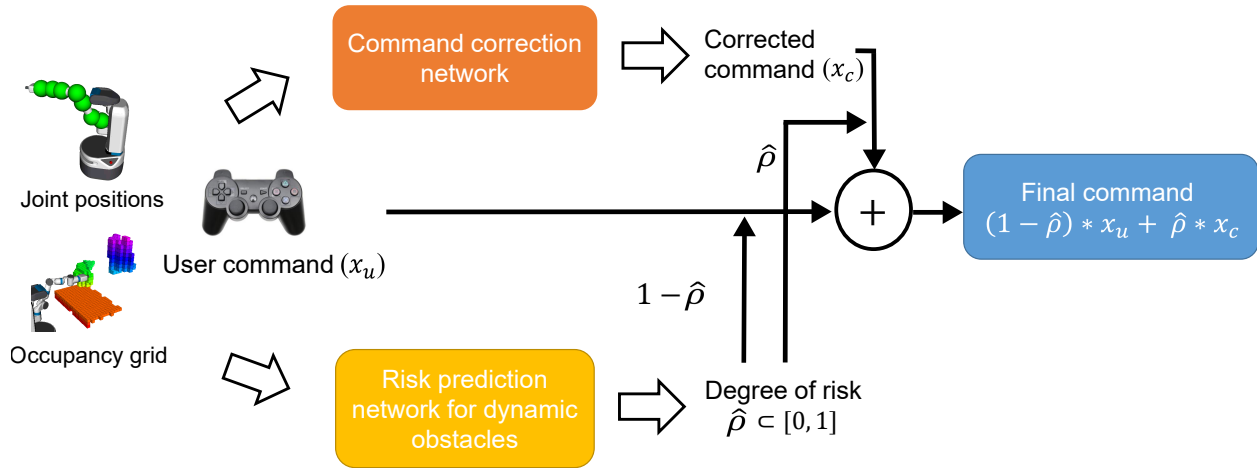


Fig. 2. This shows the system flow of our approach.

configurations for one end-effector pose due to having greater than six degrees of freedom (DoF).

Our goal is to perform safe remote manipulation in environments with dynamic obstacles. An IK solver considering collision avoidance can ensure safety for static obstacles by stopping a robot but cannot guarantee safety for dynamic obstacles. Even though a robot can be moved to a safe place through the user’s judgment, a human can make mistakes. Therefore, we need a way to protect a robot from an unrealizable command causing a collision with obstacles.

In this work, we aim to correct unrealizable user commands by predicting the risk of collision with dynamic obstacles.

B. User command correction

Since the main purpose of remote manipulation is for a robot to follow a user’s command, the command should be corrected only when it is judged to be unrealizable. We judge that the case of having the risk of collision with dynamic obstacles is highly likely to have unrealizable commands. Hence, we predict the risk of dynamic obstacles and correct an unrealizable command to move to a safe region. To do that, we use two kinds of networks. One is a risk prediction network (RPN) for dynamic obstacles, and the other is a command correction network (CCN).

Fig. 2 shows our system flow. The RPN predicts the degree of risk, $\hat{\rho} \in [0, 1]$, for dynamic obstacles from consecutive occupancy grids and joint positions. The CCN generates a corrected command, $x_c \in \mathbb{R}^6$, to move away from the obstacle by taking the current environment and robot state into account. Finally, we decide the final command, x_f , between a user command x_u and a corrected command x_c according to $\hat{\rho}$: $x_f = (1 - \hat{\rho}) * x_u + \hat{\rho} * x_c$. In short, when $\hat{\rho}$ is high, x_f is close to x_c , and when $\hat{\rho}$ is low, x_f is close to x_u . The final command x_f is delivered to an IK solver considering collision avoidance to synthesize a joint configuration; we use RCIK [5] as the IK solver.

III. CONCLUSION

In this paper, we introduced a user command correction method for safe remote manipulation in dynamic environments. We presented two networks, the risk prediction network for dynamic obstacles and the command correction network. Based on the predicted risk, our method decides the final command whether to follow a user’s command or a corrected command. In the future, we would like to advance the proposed methods and prove their robustness through various experiments.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (No. 2021R1A4A1032582).

REFERENCES

- [1] Shinji Kawatsuma, Mineo Fukushima, and Takashi Okada, “Emergency response by robots to fukushima-daiichi accident: summary and lessons learned”, *Industrial Robot: An International Journal*, 2012.
- [2] Cai Meng, Tianmiao Wang, Wusheng Chou, Sheng Luan, Yuru Zhang, and Zengmin Tian, “Remote surgery case: robot-assisted teleneurosurgery”, in *ICRA*. IEEE, 2004, vol. 1, pp. 819–823.
- [3] Jakob Thumm and Matthias Althoff, “Provably safe deep reinforcement learning for robotic manipulation in human environments”, in *ICRA*. IEEE, 2022.
- [4] Daniel Rakita, Haochen Shi, Bilge Mutlu, and Michael Gleicher, “Collisionik: A per-instant pose optimization method for generating robot motions with environment collision avoidance”, in *ICRA*. IEEE, 2021.
- [5] Mincheul Kang, Yoonki Cho, and Sung-Eui Yoon, “RCIK: Real-time collision-free inverse kinematics using a collision-cost prediction network”, *RA-L*, vol. 7, no. 1, pp. 610–617, 2021.

Fast and Robust Trajectory Generation for Cartesian Path-following Problems of Redundant Manipulators

Minsung Yoon[†], Mincheul Kang[†], Daehyung Park^{†‡} and Sung-Eui Yoon^{†‡}

Abstract—It is an important problem that quickly finds a joint trajectory so that an end-effector path of the trajectory precisely follows the Cartesian path defined in $SE(3)$. However, as the length of the considered path or the degree of freedom of the robot increases, it becomes very complicated to find a trajectory that satisfies necessary constraints such as continuity, mechanical limits, singularity, and collision avoidance. Therefore, we present a learning-based trajectory generation framework that can rapidly produce a joint trajectory while satisfying constraints with generalizability in the configuration of the external environment and the path. Our Markov decision process formulation enables our trained policy to generate trajectories with a lower constraint violation rate than the three other trajectory generation baseline methods.

I. INTRODUCTION

A path-following problem of manipulators is an important issue for real-world tasks in remote control or other domains. Given a fully constrained (i.e., 6-dimensional) path, we aim to find a configuration-space trajectory for kinematically redundant manipulators taking into consideration a variety of trajectory constraints, such as joint continuity, smoothness as well as potential collision in the environment.

Traditionally, inverse kinematics (IK) has been utilized to find a joint configuration given an end-effector pose [1]. However, the IK approach does not consider the constraints that arise when tracing the path, leading to a situation where a valid configuration no longer exists. Therefore, global search approaches have built a discrete layered graph with IK solutions computed along the path to search feasible trajectories [2]. Although these approaches are asymptotically optimal, they are quite slow since the redundant manipulator has infinite IK solutions even at one end-effector pose. As trajectory optimization has been widely adopted for generating a feasible trajectory, previous method [3] append constraints on the end-effector poses to follow the path. However, an optimized result is highly sensitive to an initial guess, i.e., initial trajectory, since these constraints cause many local minima and the local nature of the optimization method.

To balance a generation time and the trajectory quality, we amortize the online computation to satisfy the constraints by training our neural network policy offline using reinforcement learning (RL) with a variety of path-following problems and demonstrations. Our method generates the trajectory sequentially by selecting the extension direction in configuration

space based on the path and environment information starting from an initial configuration. To learn such behavior, we formulate the path-following problem as a finite-horizon Markov decision process (MDP) by defining a unified reward function composed of the task, imitation, and constraint-relevant rewards. In addition, for generalizability over diverse path-following problems (e.g., paths, start configurations, and environments), we generate diverse training environments.

We experimented with a 7 DoF Fetch robot and compared our work with the conventional IK method and a supervised learning-based method as another learning-based method. As a result, it showed faster generation time and a lower constraint violation rate with the improved null-space continuity.

II. RL-BASED TRAJECTORY GENERATION

A. Notations

Researchers often represent the path in Cartesian space as a sequence of poses $X = [x_0, x_1, \dots, x_{N-1}] \in \mathcal{X}$ evenly spaced in time, where N is the number of poses in the path, \mathcal{X} is the space of paths, and each pose is a pair of position ($\in \mathbb{R}^3$) and orientation ($\in SO(3)$). Likewise, the joint trajectory is a sequence of joint configurations $\xi = [q_0, q_1, \dots, q_{N-1}] \in \Xi$, where $q \in \mathbb{R}^d$ is a configuration of a d DoF manipulator and Ξ is the Hilbert space of joint trajectories. The d is greater than 6 in the case of redundant manipulators in $SE(3)$.

B. Formulation of MDP

We first formulate a path-conditioned MDP as $\mathcal{M}_X = \langle \mathcal{S}, \mathcal{A}, \mathcal{R}_Z, \mathcal{T}, \mathcal{Q}_0, \gamma \rangle_{X \sim \mathcal{P}(X)}$, where X is a target path sampled from a distribution of paths $\mathcal{P}(X)$, \mathcal{S} is a set of states, \mathcal{A} is a set of actions, $\mathcal{R}_Z : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is a time-varying reward function, where $Z = \{z \in \mathbb{N}_0 | 0 \leq z \leq N-1\}$ is a set of time steps, $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$ is a deterministic transition function, \mathcal{Q}_0 is a set of start configurations q_0 , and $\gamma \in [0, 1)$ is a discount factor. We sample the target paths X within a restricted operation range with the arm's length in the task space. In the case of \mathcal{Q}_0 , we sample IK solutions at the first pose x_0 in X . The policy trained on the MDP \mathcal{M}_X synthesizes a high quality trajectory ξ by sequentially expanding the trajectory in a direction that satisfies the constraints. To generalize the policy over the path and obtain a unified policy, we define a multi-path RL objective function:

$$\underset{\pi}{\text{maximize}} \mathbb{E}_{X \sim \mathcal{P}(X)} \left[\mathbb{E}_{\substack{(s_i, a_i) \sim \rho_\pi \\ q_0 \sim \mathcal{Q}_0}} \left[\sum_{i=0}^{N-1} \gamma^i \cdot \mathcal{R}_i(s_i, a_i) \right] \right], \quad (1)$$

[†]M. Yoon, M. Kang, D. Park, and S. Yoon are with the School of Computing, Korea Advanced Institute of Science and Technology, South Korea; {minsung.yoon, mincheul.kang, daehyung}@kaist.ac.kr, and sungeui@kaist.edu.

[‡]D. Park and S. Yoon are co-corresponding authors.

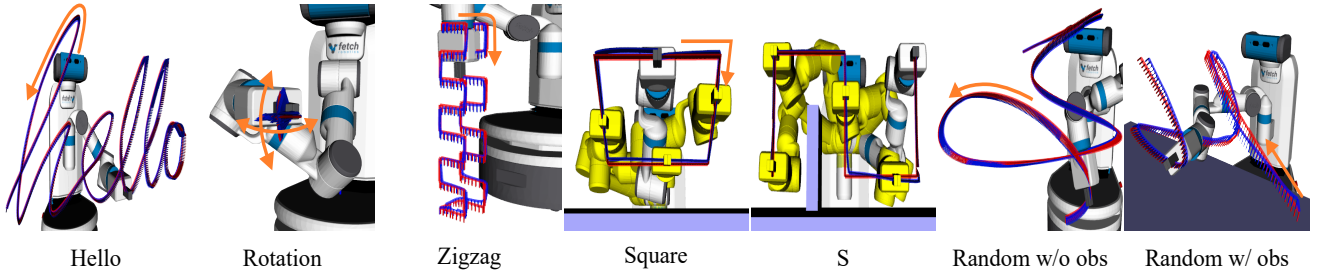


Fig. 1. Visualization of five specific and two exemplar random target paths (red lines) used in evaluations. Orange arrows indicates the progress direction of the path. Blue lines are the end-effector paths calculated from *RL-TG*'s joint trajectories via forward kinematics. In Square and S problems, the original color of the robot represents the initial configurations, and the Yellow trails indicate that the generated trajectories satisfy collision-avoidance constraints.

where ρ_π is the trajectory distribution given the deterministic transition function \mathcal{T} and a stochastic policy $\pi(a_i|s_i)$. Then, we find an optimal policy π^* where $\pi^* : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ maximizes the objective function.

1) *State and Action space*: As the state space \mathcal{S} , we consider the 3D occupancy grid for recognizing the surrounding environment, the joint values and the poses of links of the robot, and the poses up to 6 steps ahead of the target paths for near-sight behavior. We also define an action as a configuration difference, $a_i = \Delta q_i \in \mathbb{R}^d$, and $q_{i+1} = q_i + \Delta q_i$ given the deterministic \mathcal{T} . Therefore, the policy sequentially extends the trajectory for every step to compose the whole trajectory of length N .

2) *Reward formulation*: A multi-objective reward function is divided into three main terms. First, a task reward is to encourage the agent to follow the target path X . Second, an imitation reward is to make the policy learn the optimized null-space motion depicted in the demonstration. The last reward term is the constraint-related reward function that penalizes collision, joint-limit violation, singularity condition, and early termination states to satisfy constrains.

III. RESULTS

We prepared five specific ('Hello', 'Rotation', 'Zigzag', 'Square' and 'S') and one random target path benchmark set to show generalizability over external environments and paths. As baselines, *Linear* returns a linearly interpolated trajectory in joint space, *Greedy* [3] efficiently uses IK solutions, and *BC-TG* is the same as ours except learning method with supervised learning. We call our method as *RL-TG*.

Fig. 1 shows the trajectories synthesized by *RL-TG* for each exemplar benchmark problem. Fig. 2 shows the comparative analysis of the trajectory generation methods in terms of four quality metrics: a path-following pose error and a trajectory smoothness to measure a trajectory quality, a constraint violation rate considering the collision-free and joint velocity limit violation constraints, and a generation time. *RL-TG*, a learning-based function approximation method, has a slight error with the target path when looking at the end-effector path of the generated trajectory, but shows great advantages in null-space continuity (joint smoothness) and generation time, which leads to a low constraint violation rate. In Fig. 2-(d), the reason why the *Linear*'s constraint violation is zero is that the generated trajectory does not move at all since the start and end poses of the paths are the same.

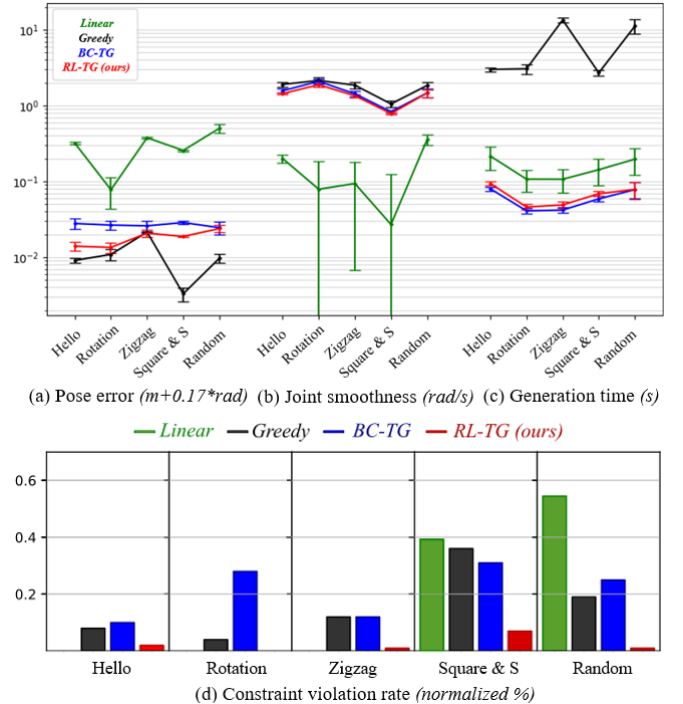


Fig. 2. Comparative analysis of the four trajectory generation methods in five types of simulated environments. The x- and y-axes are the type of benchmark problems and the performance metrics, respectively.

IV. CONCLUSION

We presented a reinforcement learning-based trajectory generation (*RL-TG*) method that quickly finds a low-cost trajectory of redundant manipulators for path-following problems. We have shown the lower constraint violation rate and fast generation time of *RL-TG* qualitatively and the generated trajectories qualitatively in simulation experiments. One future direction is to post-process the generated trajectories using trajectory optimization techniques to guarantee the generated trajectories' feasibility and to reduce the path-following error.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A4A1032582).

REFERENCES

[1] Patrick Beeson and Barrett Ames, "TRAC-IK: An open-source library for improved solving of generic inverse kinematics", in *IEEE International Conference on Humanoid Robots*. IEEE, 2015, pp. 928–935.

- [2] Rachel Holladay, Oren Salzman, and Siddhartha Srinivasa, “Minimizing task-space frechet error via efficient incremental graph search”, *RA-L*, vol. 4, no. 2, pp. 1999–2006, 2019.
- [3] Mincheul Kang, Heechan Shin, Donghyuk Kim, and Sung-Eui Yoon, “TORM: Fast and accurate trajectory optimization of redundant manipulator given an end-effector path”, in *IROS*. IEEE, 2020, pp. 9417–9424.

DOA Estimation based on Learnable TDOA Feature

Inkyu An¹ and Sung-eui Yoon²

Abstract—We propose the novel learnable time difference of arrival model to estimate direction-of-arrivals (DOAs). Different from existing TDOA features, our learnable TDOA model can consider semantic information of sound and distinguish various sound events. Our TDOA model is based on the self-attention mechanism. The self-attention mechanism is effective in learning semantic information of sound by analyzing the spatial relation of input audio.

Our TDOA model are designed to learn TDOA information of sound events effectively. We compared our approach to the prior work, i.e., the speech-oriented sound source localization task [1]. We observed a significant improvement, i.e., a 32.7 % error reduction in the mean absolute error and a 2.1 % improvement in accuracy, compared to the prior work.

I. INTRODUCTION

Sound source localization (SSL) is a fundamental problem for robot auditions. There have been many efforts based on signal-processing-based techniques to deal with the SSL problems. Although meaningful progress exists thanks to signal-processing-based approaches, many issues remain in SSL.

Deep learning (DL)-based methods have recently been presented, and they give us significant improvements. He *et al.* [1] proposed the deep neural networks for multiple speaker localization. By adding noises, e.g., fan noises of the robot, to the training dataset, their method can be robust against the noises. Adavanne *et al.* [2] proposed the source localization method for multiple sound events, e.g., alarm, speech, and footstep. Their method can detect and localize multiple sound events simultaneously. These DL-based SSL methods were trained by multiple-channel audio datasets [1], [3].

It is popular for prior DL-based methods to utilize audio features based on signal processing techniques. To localize sound source positions, the time difference of arrival (TDOA) features, e.g., generalized cross correlation-phase transform (GCC-PHAT) [4], are widely used. The combination of TDOA features given a microphone array corresponds to the specific direction-of-arrival (DOA); thus, prior DL methods estimates DOAs by considering the combination of TDOA features. However, existing TDOA features can only encode the time difference between two coherent signals, but cannot consider semantic information of sound events.

This research was supported by the MSIT, Korea, under the ITRC support program(IITP-2022-2020-0-01460) supervised by the IITP

¹Inkyu An is with School of Computing, Korea Advanced Institute of Science and Technology, Daejeon, South Korea, inkyu.an@kaist.ac.kr

²Sung-eui Yoon is with Faculty of School of Computing, Korea Advanced Institute of Science and Technology, Daejeon, South Korea sungeui@kaist.edu

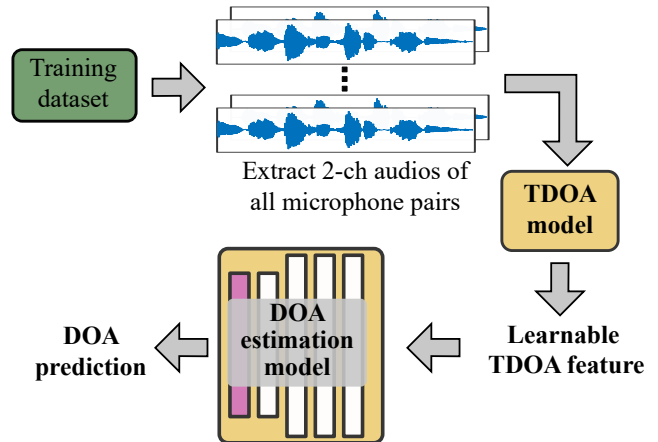


Fig. 1. The overview of our approach. 2-ch audios of microphone pairs in the dataset become an input of our approach. Our TDOA model learn to estimate TDOAs considering semantic information of sound. Our TDOA model computes multiple learnable TDOA features. Our DOA estimation model is designed to estimate DOAs from multiple learnable TDOA features. Our DOA estimation model learns the relation between the combination of multiple learnable TDOA features and the corresponding DOA. The white and purple boxes indicate the linear layer and the Sigmoid function, respectively.

We propose the novel learnable TDOA feature. Our TDOA feature, computed by our TDOA model, can learn semantic information of sound as well as time differences between two coherent audio signals. Our model can be trained by the training dataset and show better performance compared to the prior work [1]. Our model can distinguish various sound events by considering semantic information.

II. THE DOA ESTIMATION USING LEARNABLE TDOA FEATURE

The overview of our approach is shown in Fig. 1. We first extract every 2-ch audio of all microphone pairs from the training dataset. Extracted 2-ch audios enters to our TDOA model; our TDOA model can estimate TDOAs of two coherent sound signals considering semantic information. After our TDOA model computes the learnable TDOA features, we can estimate DOAs from the learnable TDOA features.

Learnable TDOA feature. Our TDOA model is effective to estimate TDOAs by consider semantic information. Our TDOA model is designed based on the self-attention mechanism [5]. The self-attention mechanism is a powerful approach to analyze spatial relations of input [6], [7]. For example, given an image containing a person, the self-attention mechanism can learn the spatial relation between each part of the person, e.g., head, body, and legs.

TABLE I

THE ACCURACY OF DOA ESTIMATIONS OF OURS AND THE PRIOR WORK.

	MAE (\downarrow)	ACC (\uparrow)
MLP-GCC [1]	4.61 degree	91.91 %
Ours	3.1 degree	93.9 %

The self-attention mechanism is also useful to learn semantic information of sound. The audio input contains consecutive sound signals capturing various sound events, e.g., speech and footstep. Each sound event has different spatial relations; for example, speech is a sequence of voices of a person, but footstep is a sequence of the sound of a person’s foot tapping the floor. Our approach utilize the self-attention mechanism to learn spatial relations in the time domain and, thus, can distinguish different sound events.

The input of the TDOA model is 2-ch audios of all microphone pairs; thus, there should be multiple learnable TDOA features given all microphone pairs. The multiple learnable TDOA features go to the next step: estimating DOAs.

The DOA estimation model. The combination of TDOAs of all pairs given the microphone array corresponds to the specific DOAs; thus, we can estimate DOAs from the combination of learnable TDOA features. Our DOA estimation model is designed to learn those relations between the combination of TDOA features and the corresponding DOA.

Our DOA estimation model consists of four linear layers, and the Sigmoid function computes the DOA predictions. We utilize the binary cross-entropy loss between the DOA predictions and DOA labels to train our models.

III. RESULT AND DISCUSSION

We tested our approach in the speech-oriented SSL, i.e., estimating DOAs only of speech sounds. We compared our approach to the prior works [1] utilizing the existing TDOA feature, e.g., GCC-PHAT [4]. By comparing to prior works, we want to show the effectiveness of our learnable TDOA feature. We utilize the SSLR dataset [1] recorded from various conversations with additional noises. The SSLR dataset is recorded by the 4-ch circular microphone array; thus, there exist six microphone pairs in four microphones.

We utilize the evaluation metric proposed by [1], consisting of MAE and ACC; they are the mean absolute error and the accuracy of correct predictions, respectively.

We verify that our approach shows better accuracy for both metrics than the prior work in Table. I. We observe that our approach gives us a significant improvement, i.e., a 32.7 % error reduction in MAE and a 2.1 % improvement in ACC, compared to the prior work.

Those results show that our learnable TDOA feature is useful for localizing sound sources. Our learnable TDOA feature can be more helpful than existing TDOA features, e.g., GCC-PHAT. Moreover, we verify that our learnable TDOA feature efficiently considers semantic information. Considering semantic information is important to localize various speech in SSLR while ignoring additional noises.

REFERENCES

- [1] Weipeng He, Petr Motlicek, and Jean-Marc Odobez, “Deep neural networks for multiple speaker detection and localization”, in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 74–79.
- [2] Sharath Adavanne, Archontis Politis, Joonas Nikunen, and Tuomas Virtanen, “Sound event localization and detection of overlapping sources using convolutional recurrent neural networks”, *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 1, pp. 34–48, 2018.
- [3] Archontis Politis, Sharath Adavanne, Daniel Krause, Antoine Deleforge, Prerak Srivastava, and Tuomas Virtanen, “A dataset of dynamic reverberant sound scenes with directional interferers for sound event localization and detection”, *arXiv preprint arXiv:2106.06999*, 2021.
- [4] C. Knapp and G. Carter, “The generalized correlation method for estimation of time delay”, *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327.
- [5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin, “Attention is all you need”, *Advances in neural information processing systems*, vol. 30, 2017.
- [6] Anurag Arnab, Mostafa Dehghani, Georg Heigold, Chen Sun, Mario Lučić, and Cordelia Schmid, “Vivit: A video vision transformer”, in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 6836–6846.
- [7] Kai Han, An Xiao, Enhua Wu, Jianyuan Guo, Chunjing Xu, and Yunhe Wang, “Transformer in transformer”, *arXiv preprint arXiv:2103.00112*, 2021.

Judging by the Look: The Impact of Robot Gaze Strategies on Human Cooperation

Di Fu^{†*}, Fares Abawi^{†*}, Erik Strahl[†] and Stefan Wermter[†]

Abstract—Human eye gaze plays an important role in delivering information, communicating intent, and understanding others’ mental states. Previous research shows that a robot’s gaze can also affect humans’ decision-making and strategy during an interaction. However, limited studies have trained humanoid robots on gaze-based data in human-robot interaction scenarios. Considering gaze impacts the naturalness of social exchanges and alters the decision process of an observer, it should be regarded as a crucial component in human-robot interaction. To investigate the impact of robot gaze on humans, we propose an embodied neural model for performing human-like gaze shifts. This is achieved by extending a social attention model and training it on eye-tracking data, collected by watching humans playing a game. We will compare human behavioral performances in the presence of a robot adopting different gaze strategies in a human-human cooperation game.

I. INTRODUCTION

Eye contact plays an important role in interpersonal communication. A recent study shows eye contact can increase the synchronization between humans’ brains [1]. Maintaining eye contact can deliver information, allowing for the inference of intent, and understanding others’ mental states. Likewise, robot gaze can influence human decision-making during gameplay by lowering reaction time and increasing the cognitive effort [2]. The authors, however, design a robot gaze strategy following heuristic findings rather than simulating human eye movements, resulting in less realistic gaze shifts by the robot. Another study introduces a social gaze-control system to simulate visual human attention [3]. However, there is no direct human-robot interaction in the experimental scenario. Moreover, the impact of robot gaze on human decision-making is not studied.

Overall, previous work suffers two crucial limitations: 1) Gaze simulation is not based on eye-tracking data; 2) Due to the disregard of human eye movements in designing robot gaze strategies, experiments do not examine the influence of such effects on humans. To address the aforementioned gaps, we propose two research goals: 1) Simulate human gaze behaviors on robots based on eye-tracking data; 2) Explore the influence of a robot’s gaze interaction on the performances of humans in a cooperation game.

II. INTERACTION SCENARIO, TASKS, AND PROCEDURE

Our experiments will be performed in three separate tasks conducted in sequence as illustrated in Figure 1.

A. Task 1: Human-human cooperation game with an inactive robot

In task 1, participants will be randomly matched in pairs to play multiple rounds of a human-human cooperation game. We will use the iCub¹ robot in our experiments. The robot will display facial expressions in a Wizard-of-Oz setup without performing gaze or head movements. During the task, a pair of participants will sit around a table facing each other and play a game. The iCub robot will act as the instructor by asking one of the participants to place a particular shape in its corresponding hole on a shape sorter, e.g. “*place the cylinder in the round hole*”. One participant will play the role of an actor, placing his or her hands in a box occluding the available objects and containers. The other participant will assume the role of a guide by helping the actor place the right object in the designated container through speech and gestures. After each round, the iCub robot will change its facial expressions. The participants will be asked to guess the intention behind the different facial expressions during the experiment. In doing so, the participants would distribute their attention between the task at hand and the iCub robot. The guiding participant will be asked to push a button once the round is completed to measure the response time. There will be 20 rounds in total. Each pair of participants will flip roles as actors and guides after every 5 rounds by switching seats. The 20-round response time will be summed up to be the total cooperation time for each pair. After finishing the cooperation game, the participants will fill in the Godspeed questionnaire [4] to rate their impression of the iCub robot. The entire game session will be recorded using the iCub’s cameras and binaural microphones.

B. Task 2: Social attention model training on human eye-tracking data

In task 2, we will collect human eye-tracking data by recruiting participants to watch the videos recorded in task 1. We will then train a social attention model based on this eye-tracking data. The videos will be displayed on a large curved screen with stereo audio playback. The participants’ eye movements will be recorded under the free-viewing condition

*Di Fu and Fares Abawi contributed equally to this work.

[†]Di Fu, Fares Abawi, Erik Strahl, and Stefan Wermter are with Knowledge Technology, Department of Informatics, University of Hamburg, 22527 Hamburg, Germany {di.fu, fares.abawi, erik.strahl, stefan.wermter}@uni-hamburg.de

¹iCub: <https://icub.iit.it/>



Fig. 1. Three tasks comprising our experimental setup starting with A) task 1 in which the robot remains static and records a pair of participants playing a cooperation game with a shape sorter (top left), followed by B) task 2 where the head and eyes of a participant watching the gameplay are recorded using an eye tracker (top left). In C) task 3 the robot performs movements based on either strategy while participants play the game described in task 1.

while wearing a Pupil Core eye tracker². The camera view in task 1 is from the iCub’s perspective, therefore, the robot is not visible in task 2. A saliency prediction model [5] with crossmodal social cue integration will be trained on the collected human eye-tracking data. The model will be extended with a mapping scheme for associating perceivable stimuli within the auditory and visual fields. Such a model requires binaural perception to localize sound sources when visual stimuli do not lie within the visual field. Our model will, therefore, combine priority maps arriving from different modalities to form a master map for attending to elements with high conspicuity, regardless of their visibility.

C. Task 3: Human-human cooperation game with human-robot joint and mutual gaze interaction

The social attention model trained in task 2 will be mounted on the iCub robot. In this task, there will be two strategies guiding the robot’s gaze behavior - interactive and random. The interactive strategy will follow the predictions of our social attention model trained on the eye gaze data acquired in task 2. This strategy will likely increase the robot’s joint and mutual eye gaze with the participants. The random strategy model will, however, perform gaze shifts based on statistical assumptions regarding fixation time and saccadic movements. This model is agnostic to the stimuli, allowing for robot behavior similar to the typical human without perceiving the environment. Pairs of participants will be recruited to play the same game described in task 1, for a total of 40 rounds. As opposed to task 1, the iCub will perform gaze movements based on either strategy. The participants’ response time will be recorded given their exposure to the iCub following the interactive and random gaze strategies for 20 rounds per strategy. The condition (strategy) ordering will be shuffled to avoid order bias.

III. DATA ANALYSES AND EXPECTED RESULTS

The sum and the average reaction time for each pair of participants will be calculated. One-way ANOVA will be measured on these statistics for the three robot conditions (inactive/interactive/random) to detect the impact of different human-robot gaze interactions on human cooperation. Ideally, if the pairs remain unchanged between tasks 3 and 1,

an independent t-test could be implemented by comparing the test-retest differences between the interactive and random groups. The inactive condition will be taken as a first test baseline. For the analyses of participants’ ratings on the Godspeed questionnaire, one-way ANOVA will be measured for the sum of different dimensions across three groups.

We hypothesize that the social attention model can integrate multiple crossmodal cues and predict social saliency based on human data. Moreover, we expect that the interactive gaze strategy could increase human-human cooperation by decreasing the reaction time during gameplay compared to the other conditions.

IV. CONCLUSIONS

We propose to train a social attention model on human gaze behaviors in a crossmodal environment. The model will be deployed on the iCub robot as an embodiment platform. Human decision-making and cooperation behaviors will be studied under different robot gaze conditions to explore the impact of joint and mutual attention in human-human and human-robot interaction. Our work combines behavioral findings and computational modeling to bridge the gap between cognitive simulation and human-robot interaction.

ACKNOWLEDGMENT

The current work is supported and funded by the German Research Foundation DFG under project CML (TRR 169).

REFERENCES

- [1] C. D. B. Luft, I. Zioga, A. Giannopoulos, G. Di Bona, N. Binetti, A. Civilini, V. Latora, and I. Mareschal, “Social synchronization of brain activity increases during eye-contact,” *Communications Biology*, vol. 5, no. 1, pp. 1–15, 2022.
- [2] M. Belkaid, K. Kompatsiari, D. De Tommaso, I. Zablit, and A. Wykowska, “Mutual gaze with a robot affects human neural activity and delays decision-making processes,” *Science Robotics*, vol. 6, no. 58, p. eabc5044, 2021.
- [3] A. Zaraki, D. Mazzei, M. Giuliani, and D. De Rossi, “Designing and evaluating a social gaze-control system for a humanoid robot,” *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 157–168, 2014.
- [4] C. Bartneck, E. Croft, and D. Kulic, “Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots,” *International Journal of Social Robotics*, vol. 1, pp. 71–81, 2009.
- [5] F. Abawi, T. Weber, and S. Wermter, “GASP: Gated Attention for Saliency Prediction,” in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, IJCAI Organization. International Joint Conferences on Artificial Intelligence, 2021, pp. 584–591.

²Pupil Core: <https://pupil-labs.com/products/core/>

Virtual Reality Framework for Better Human-Robot Collaboration and Mutual Understanding

Maciej K Wozniak* and Patric Jensfelt

Abstract—Humans interact with robotic systems on a daily basis. User-friendly and efficient interfaces connecting us with these systems are critical for efficient collaboration and a good user experience. In the latest machine learning developments, many robotic platforms have used deep learning models to understand the environment and surroundings better. However, what a robot senses and how it takes decisions are usually hidden from the user. It is believed that soon we will be able to work side-by-side with these machines in a connected, collaborative space. Thus, it is essential to understand the robot and easily reason with it about the state of the environment or how it wants to execute a particular task. This work presents a virtual reality (VR) framework for human-robot collaboration, focused on improving communication and understanding between humans and robots.

I. INTRODUCTION

Understanding the environment is one of the crucial tasks most robots must perform. The goal is not to hard-code each possible action, but to allow the robot to reason about the surroundings and learn how to move and act within a changing environment. Even though the goal may vary from delivering a package [1] to cleaning an apartment [2], comprehensive knowledge of the environment is required to complete the task. Another important factor is mutual trust and understanding between the user and the robot, so both can reason and collaborate safely and efficiently. Therefore, it is critical to see how the robot perceives the environment.

An interaction between a user and a robot may take various forms. An example of a commonly used one is verbal communication [3]. A user may ask a robot to move an object (e.g., a tennis ball) in a specific space (e.g., a room). To accomplish this (i.e., move the ball from point A to B), the robot has to understand the sentence (e.g., using natural language processing), execute the action and update its understanding of the environment. However, many errors may arise from an inaccurate sentence, e.g., *remove the ball from the table*; the robot may not understand *which ball should be moved* or *where it should be placed*. Such confusion may cause the whole system to fail [3].

To overcome these challenges, we propose an intuitive VR interface where a user can easily communicate with the robot. We create a 3D virtual representation of the real world, which the user sees and interacts with in VR. An interactive virtual interface shows what the robot understood about the environment in real-time so that the user can correct the robot's reasoning and actions if necessary (described in detail

in Section VI). This approach creates a clear understanding of the task and helps to eliminate possible errors. Now, the user can grab an object in the virtual environment and place it somewhere else as a way to instruct the robot about what to do in the real world. Because the virtual environment was created based on the output of the robot's perception module, the user understands what the robot knows about its workspace and can seamlessly communicate to the robot what actions it has to perform.

Additionally, the trajectory and movement of a robot may not always fit the users' preferences. For example, a robot is usually biased to choose the shortest path; however, humans' preferences of the robot's actions may differ. They can feel afraid and uncomfortable being side-by-side with the robot, not knowing what it will do if they are not used to being around a robot. Consequently, they may prefer to see what the robot intends to do beforehand so that they can approve or disapprove its actions and potentially modify them. If users could do that, they would most likely be less hesitant to work and share a common space with it.

Building trust and creating seamless human-robot interaction was the primary motivation for our work. Seeing precisely what a robot understands and what actions it will take brings us to another level of trust, which often cannot be achieved even in human-human interaction.

Our main contribution is a virtual reality framework for human-robot collaboration. Our tool serves the purpose of building users' trust through understanding how the robot perceives its surroundings and reasoning with it about its actions. Its main features are the ability to visualize what the robot understands about the environment, interact with it and alter its actions through an immersive VR user interface (UI).

We tested the framework within a virtual environment using a simulation engine and VR headset. We emphasize that an immersive VR UI elevates the interaction and communication between the user and the robot. VR makes it easy and intuitive to understand and perform spatial tasks such as 3D trajectory modification. Our work promotes the use of VR for human-robot collaboration and shows yet another application of VR in robotics-oriented projects.

II. RELATED WORK

In this Section, we will review some of the recent projects focused on improving human-robot collaboration using virtual/augmented reality (VR/AR) devices. Reviewed articles also show how important it is for users to be able to reason,

*This research was supported by Wallenberg AI, Autonomous Systems and Software Program (WASP). Contact info: {maciejw,patric}@kth.se, KTH Royal Institute of Technology, Stockholm 10044, Sweden

understand, and discuss with the robot about its actions and intentions.

In recent years, researchers realized that VR elevates the interaction and communication between the user and the robot, making it an immersive experience. Szafir and Szafir [4] focused on data visualization between humans and robots, showing that data visualization is a fundamental aspect of good collaboration and mutual understanding between humans and robots. Other approaches focused on creating VR/AR frameworks for Human-Robot Collaboration (HRC). Mara et al. created Cobot [5] – an educational HRC platform where participants can play interactive mini-games with robots and work together on different tasks. Kennel-Maushart et al. [6] developed a tool that enables the user to manipulate robotic arms in the *real world* by applying force to them in the virtual reality setup. It is vital that when humans and robots work in the same space, robots can correctly estimate humans' locations and poses. VR headsets and controllers can greatly facilitate that task. The robot can easily track human movements using the sensors in the controllers and the headset. At the same time, the environment can be augmented and presented to the user in a different form via the headset, making the interaction more immersive and interesting [7].

These projects promote the use of VR for human-robot interaction and educate humans about robots; however, they do not focus on executing practical tasks or improving how humans interact with robots. On the other hand, our framework focuses on improving the interactions and understanding a robot's intentions for practical tasks.

Other researchers applied VR to projects in the industrial setting, creating and training specific scenarios in the virtual environment before performing them in the real world. An example could be the assembly process [8], collaborative tasks performed together with the robot [9], or visualizing safe space for physical assembly workers and robots [10].

These methods focus primarily on personnel training for manufacturing industry. Additionally, they are made in a fixed environment, whereas our framework can operate in changing environments and be used for various tasks.

Moreover, many scientists used VR to create and conduct *digital twin* experiments. A *digital twin* is a virtual representation of a robot. Modern physics engines can imitate reality in great detail, allowing a digital twin to be an accurate *test-bed* for real-world applications [11]. Many scientists, such as Kuts et al. [12] used VR in the development of a framework aimed to bridge the gap between real-world and simulation-based industrial robots. Others also showed various applications and benefits of using VR interfaces for digital twin projects, such as improved factory safety or workers' training [13].

Described projects focus on visualization and representation of a robot's action; however, they do not provide an easy-to-use interface for reasoning and correcting the robots which operate in a fixed environment. On the other hand, our work is focused on immersive and easy-to-use interface in which we can modify robot's actions. Additionally, our

perception module allows a robot to function in a changing environment.

Finally, VR/AR has started to be commonly used in robot control and teleoperation tasks [14]–[16]. Articles by Ostatin [17] and Togias [18] showed that VR is perfect for planning the trajectory of the robotic arm, allowing the users to easily plan how the robot should move. Chandan et al. [19] developed a teleoperation framework that can be used to visualize the states, intentions, and future trajectories of robots. Xu et al. [20] efficiently visualize the state of the robot's end effector in VR, allowing the user to give orders by changing the end-effector position and orientation in the VR UI.

These projects focus on planning robots' trajectories, visualizing robots' intentions, and steering the robot; however, they do not focus on the explainability aspect of what robots understand about the environment. Their visualization and manipulation are limited to either only choosing the final position of the robot's end-effector or adding *no-go* zones. Additionally, none of these projects allow the user to test their solution before deploying them to the real world. On the other hand, our framework allows the user to see how the robot perceives the environment, modify every step of its action, and verify proposed actions and trajectories before deploying them to the real world.

III. PROBLEM FORMULATION

The projects described in [Section II](#) are a significant contribution to the research community. However, we still see missing parts that are addressed in our method. A framework for human-robot collaboration should emphasize the importance of understanding a robot's intentions and its perception of the working space, something we did not explicitly see in other solutions.

A framework interface should help reduce potential errors and misunderstandings, as well as increase the level of trust between users and robots, which is essential for long-lasting collaboration. Users should be able to interact with robots without explicit programming knowledge. An interface should be immersive, simple, and intuitive to use. It should allow users to quickly test different scenarios, visualize and modify the robot's intentions (e.g., the trajectory), and effortlessly deploy the final solutions to the real world.

To sum up, a fully functional solution should enable users to:

- help the robot to avoid collision with different obstacles by modifying its path.
- incorporate the preferences about the movement of the robot.
- see what robot does and what it does not understand about the environment.
- assist it with difficult and demanding tasks where it is more likely to fail.
- give tasks to the robot.

IV. PROPOSED FRAMEWORK ARCHITECTURE

In this Section, we describe the system architecture (shown in Fig. 1). Our approach takes an important step toward a final solution for the HRC framework, providing the user with a transparent representation of how the robot perceives its surroundings and the ability to visualize and modify its intentions. The user can test different scenarios and solutions in the virtual environment before deploying them to the real world. Our framework enables the user to interact with the robot throughout the immersive user interface, giving the robot clear commands, e.g., asking it to pick and move the objects by rearranging them in the VR. Finally, we provide the functionality for visualizing and correcting the robot’s trajectory by simply grabbing and moving trajectory waypoints (intermediate points on the robot’s path).

A. VR environment

We use the Unity game engine¹ to create a VR environment and test it using the Oculus Quest 2 headsets². However, the framework can be built and run on another headset by changing the target device in Unity settings. In the rest of the text, we refer to the VR part of our framework as Virtual Reality User Interface (VR UI).

B. Environment mapping and understanding

In order to understand the environment, we used an RGB camera to collect data. The images are fed into a deep learning model (such as Detectron2 [21]) for scene segmentation and object detection to obtain segmentation masks and bounding boxes of the detected objects. This information will be later used to create and update the state of the environment.

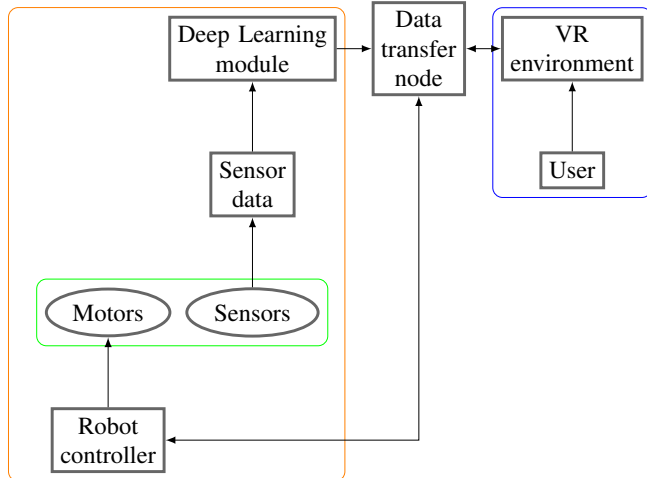


Fig. 1: Proposed system architecture. The elements within the orange box correspond to nodes and parts connected with the robot’s motion and perception, whereas the ones in the green box are explicitly corresponding to its hardware. Nodes inside the blue box are connected to the user and VR setting. The robot’s sensors collect data from the environment.

C. Data transfer

The first naive solution would be to transfer the data collected by the robot directly to the VR UI and show it to the user. However, such an approach creates various bottlenecks. It is computationally expensive to receive and render a high-quality environment in real-time. Moreover, if we use RGBD sensors or LiDAR instead of a camera, we would obtain point cloud data which is often incomplete. Thus, even though a robot understands from an image or a point cloud that an object on the table is a cup, it collects only the points from one side of the cup. Therefore, the model shown to the user in the VR would be incomplete and have multiple imperfections, such as numerous missing points. One can argue that we can solve it with shape completion or overlapping masks of the 3D models onto classified objects with those missing points. However, the lower quality of the interface could potentially worsen the overall user experience, making them more hesitant to use the tool. Additionally, such a solution would require much higher data transfer capacity or computational resources.

In order to facilitate the exchange of information, improve the user experience, and minimize the necessary bandwidth, we only transfer output of a detection network between the robot and the virtual environment. When the robot detects and classifies an object, it sends its class, location, and estimated size to the VR UI application. In the virtual environment, we have multiple prefabs (3D models of the objects we built into the project) corresponding to the detected classes, and we can quickly create a 3D representation of the room from the received message. Similarly, by reorganizing or pointing at the objects in the VR environment, we can send the request to the robot to change the position of that object in the real world. Seeing these objects and interacting with them in the VR UI help us to grasp a better understanding of how the robot perceives its environment.

V. EXPERIMENTS

The current experimental setup is shown in Fig. 2. We based our experiments on Niryo One robotic arm³; however, any robotic arm can be loaded into our framework using URDF and mesh files⁴. The environment is set up as follows. A robotic arm is located on the table with various objects placed on top. The camera is located above the table, and its output is sent to the ROS⁴ node that runs the instance detection and classification algorithm. The information obtained by the DL model is then sent to the user interface on the VR headset and the objects are created in the exact location as they were detected and classified.

Our work focuses on the VR UI part of the system, therefore, to facilitate the experiments, we created a Unity simulation corresponding to the *real-world environment*. This approach enables us to easily generate the data necessary to train the deep learning model and run the experiments in a highly controlled environment. For the machine learning

¹<https://unity.com/>

²<https://store.facebook.com/se/quest/>

³All the prefabs are available in the Unity Robotics Hub

⁴<http://wiki.ros.org/Documentation>

part, responsible for object detection and scene segmentation, we chose *Detectron2* with Faster R-CNN [22] pre-trained on COCO dataset [23]. We retrained it on the data generated in Unity using open-source prefabs (3D models such as water bottles or flashlights)⁵. The proposed framework and datasets are open-source and can be freely downloaded and tested⁶.

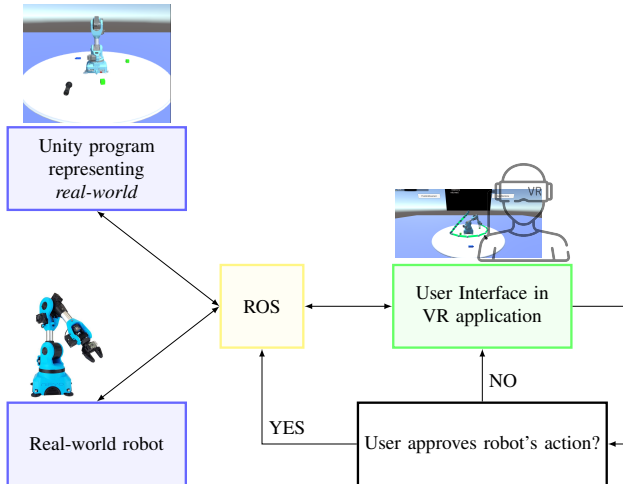


Fig. 2: Representation of experimental setup.

A. Environment understanding

In the first step, the virtual environment is created and the objects, detected and classified using the camera output from the real world, are generated in the VR UI. The user understands how the robot perceives its workspace because the interface reflects the output of the robot's perception module. If the robot's operator can only see the camera output (as in Fig. 3b) without feedback from the robot's perception module, the whole interaction may fail. A straightforward example could be a user asking a robot to pick up a tennis ball from the table. The robot may not have detected the object (e.g., a tennis ball) that the operator asks it to pick up (e.g., because the network was not trained on tennis balls Fig. 3a). Consequently, the robot will not understand that there is an object to pick up. That may lead to confusion and frustration from the users' side since they will not understand why the robot fails. Our method allows the user to see how the robot perceives the world and quickly identify such issues. The user can only interact with the objects that were correctly detected by the robot, as shown in Fig. 3c.

B. Action verification and safety

In the VR UI, objects can be selected by pointing at them with the controller as a laser pointer. Instead of testing the solution beforehand, one could quickly send such a command to a real-world robot. However, there may be potential flaws in the robot's trajectory. A simple scenario is an obstacle on the robot's course as a bottle in Fig. 4a. If we were to deploy the robot's action to the real world, we could potentially

⁵Our framework, training data and DL model are publicly available for download and testing <https://github.com/maxiuw/pickandplace>

damage the robot. Instead, our framework enables users to verify and approve the robot's trajectory before deploying it into the real world. The proposed trajectory is generated and shown in VR. Additionally, a waypoint is generated for every timestep t (value defined by the user). The user can modify the frequency and appearance of the waypoints. We use the robot's end-effector Fig. 4b or a sphere Fig. 4c to show the waypoints, but any other 3D model can be assigned to do so. Now, the user can move around and quickly see the 3D trajectory from a different perspective, which would be more complex to do using a 2D screen.

C. Trajectory modification

When we send a request to the robot, it executes the task in the VR or reports that it is not able to make a move (e.g., the target is beyond its reach). As shown in Fig. 5a, the users can see that the initial trajectory would cause a collision with another object. In that case, the proposed trajectory can be disapproved and modified. In the VR UI, the user can move the waypoints so that the robot would take a different path and avoid collision with the obstacle, as shown in Fig. 5b. The user verifies a new trajectory and the robot can execute the task following a new path (Fig. 5c). Users can examine whether now the robot executed the moves as expected and, if necessary, repeat the correction process multiple times. As we can observe in Fig. 5d, the robot does not exactly follow the new trajectory. To provide fluid motion, we minimize the number of waypoints the robot has to pass. We only include the starting and final points and the waypoints which positions were changed. That allows us to achieve a smoother transition between poses while still avoiding the obstacles⁶. Finally, when the trajectory is approved, it can be deployed and executed by the robot in the real world.

These tests allow us to eliminate potential malfunctions or hazardous situations. Additionally, we can visualize what the robot understands about the environment and its intentions in real-time. An option to review the robot's actions is essential, especially with the growing demand for home robotics. Robots are not anymore expected to work alone but to share space and collaborate with humans. Our tool will help humans feel safe and comfortable while working side-by-side with the robot.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a VR-based HRC framework and showed its capabilities in a number of tasks. In Section I, Section II, and Section III, we stated the problem of clear understanding between the user and the robot and presented existing solutions similar to ours, however, most of them lacked explainability of robot's intentions and environment understanding as well as the ability to modify actions proposed by the robot. Inspired by that, we created a HRC framework described in Section IV. In Section V, we showed how we created a representation of the real world in VR UI. Next, we presented how users can easily visualize and

⁶A short <https://youtu.be/hs3DXQhG8Ys> video showing an entire interaction with a robot.

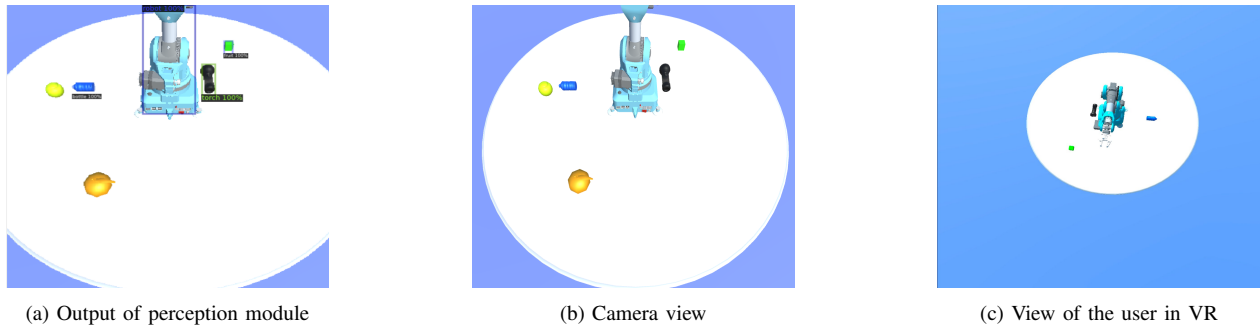


Fig. 3: Potential issues caused by robot's perception flaws.

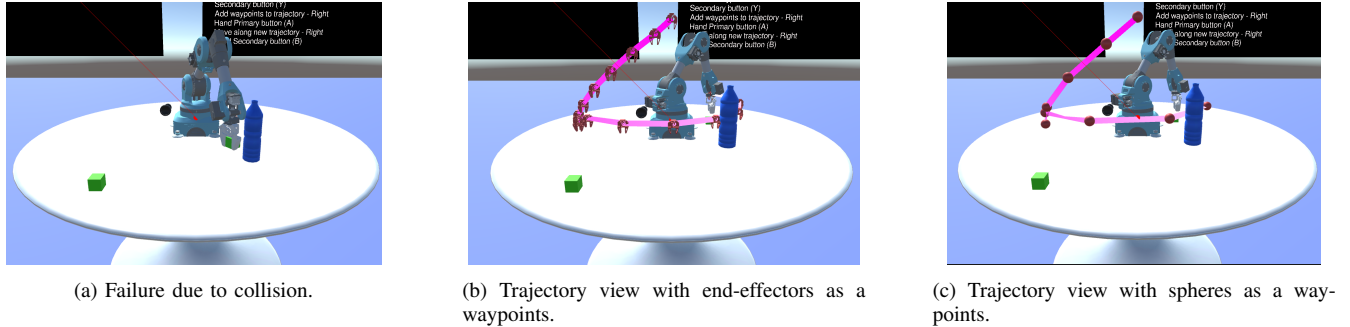


Fig. 4: Collision caused by the the obstacle on the robot trajectory. We can see that if we deploy task into the real world, without verifying it before hand, the robot may fail to complete the task (Figure a).

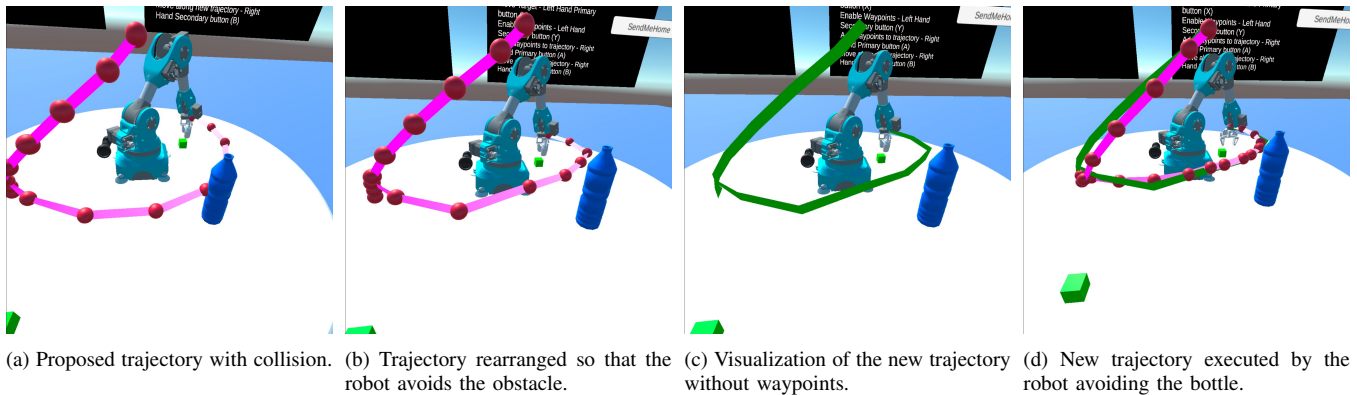


Fig. 5: Example of rearranging robots trajectory.

modify the robot's actions (e.g., to avoid obstacles) and deploy it back to the real world. Our immersive VR UI confirmed that VR is an excellent tool for interacting and collaborating with the robot. We emphasized the importance of understanding the robot's actions and abilities. While we run the experiments on a *pick-and-place example*, interacting with the robot can teach the user which tasks the robot has problems with. These aspects are fundamental, especially nowadays, when mobile robots are slowly moving to our offices [24] or homes [25].

There are still remaining challenges that we would like to tackle. First, it would be interesting to test our solution in AR setup instead of VR. This approach would allow users to see what is around them in the real world while interacting with the robot without the need to move to a fully immersive

environment. Since the project was developed in Unity, it is simple to change the target device for the software; therefore, building it on an AR headset should not be an issue.

Additionally, we would like to conduct a user study on a larger group of people to get feedback on the methods we developed and adjust our approach for the interaction to be intuitive and user-friendly. Such a study will aim to show that knowing what actions the robot will take and how it perceives the environment helps the user to gain trust in the robot's abilities.

To sum up, this project presents the framework's ability for HRC and highlights the importance of common human-robot understanding.

REFERENCES

- [1] M. Thiel, M. Grote, M. Schrick, and S. Tjaden, "Integrierte mobilität und automatisierung-transportroboter unterwegs im tabula shuttle/integrated mobility and automation-transport robots on the road in the tabula shuttle," *Automobiltechnische Zeitschrift*, vol. 124, no. 4, 2022.
- [2] K. Doelling, J. Shin, and D. O. Popa, "Service robotics for the home: A state of the art review," in *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*, 2014, pp. 1–8.
- [3] J. Peltason, N. Riether, B. Wrede, and I. Lütkebohle, "Talking with robots about objects: A system-level evaluation in hri," in *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2012, pp. 479–486.
- [4] D. Szafrir and D. A. Szafrir, "Connecting human-robot interaction and data visualization," in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 2021, pp. 281–292.
- [5] M. Mara, K. Meyer, M. Heiml, *et al.*, "Cobot studio vr: A virtual reality game environment for transdisciplinary research on interpretability and trust in human-robot collaboration," 2021.
- [6] F. Kennel-Maushart, R. Poranne, and S. Coros, "Multi-arm payload manipulation via mixed reality,"
- [7] B. Yang, X. Xie, G. Habibi, and J. R. Smith, "Competitive physical human-robot game play," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 2021, pp. 242–246.
- [8] J. Guzzi, G. Abbate, A. Paolillo, and A. Giusti, "Interacting with a conveyor belt in virtual reality using pointing gestures," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '22, Sapporo, Hokkaido, Japan: IEEE Press, 2022, pp. 1194–1195.
- [9] V. Ortenzi, M. Filipovica, D. Abdulkarim, *et al.*, "Robot, pass me the tool: Handle visibility facilitates task-oriented handovers," in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2022, pp. 1–9.
- [10] J. Zubizarreta, I. Aguinaga, and A. Amundarain, "A framework for augmented reality guidance in industry," *The International Journal of Advanced Manufacturing Technology*, vol. 102, no. 9, pp. 4095–4108, 2019.
- [11] A. A. Malik and A. Brem, "Digital twins for collaborative robots: A case study in human-robot interaction," *Robotics and Computer-Integrated Manufacturing*, vol. 68, p. 102092, 2021.
- [12] V. Kuts, T. Otto, T. Tähemaa, and Y. Bondarenko, "Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality," *Journal of Machine Engineering*, vol. 19, 2019.
- [13] X. Wang, C. J. Liang, C. C. Menassa, and V. R. Kamat, "Interactive and immersive process-level digital twin for collaborative human–robot construction work," *Journal of Computing in Civil Engineering*, vol. 35, no. 6, p. 04021023, 2021.
- [14] F. Nenna and L. Gamberini, "The influence of gaming experience, gender and other individual factors on robot teleoperations in vr," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 2022, pp. 945–949.
- [15] A. Naceri, D. Mazzanti, J. Bimbo, *et al.*, "Towards a virtual reality interface for remote robotic teleoperation," in *2019 19th International Conference on Advanced Robotics (ICAR)*, 2019, pp. 284–289.
- [16] E. Rosen, D. Whitney, E. Phillips, D. Ullman, and S. Tellex, "Testing robot teleoperation using a virtual reality interface with ros reality," in *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*, 2018, pp. 1–4.
- [17] M. Ostanin and A. Klimchik, "Interactive robot programming using mixed reality," *IFAC-PapersOnLine*, vol. 51, no. 22, pp. 50–55, 2018.
- [18] T. Togiass, C. Gkournelos, P. Angelakis, G. Michalos, and S. Makris, "Virtual reality environment for industrial robot control and path design," *Procedia CIRP*, vol. 100, pp. 133–138, 2021.
- [19] K. Chandan, V. Kudalkar, X. Li, and S. Zhang, "Arroch: Augmented reality for robots collaborating with a human," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2021, pp. 3787–3793.
- [20] S. Xu, S. Moore, and A. Cosgun, "Shared-control robotic manipulation in virtual reality," *arXiv preprint arXiv:2205.10564*, 2022.
- [21] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, *Detectron2*, <https://github.com/facebookresearch/detectron2>, 2019.
- [22] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.
- [23] T.-Y. Lin, M. Maire, S. Belongie, *et al.*, "Microsoft coco: Common objects in context," in *European conference on computer vision*, Springer, 2014, pp. 740–755.
- [24] M. Hanheide, M. Göbelbecker, G. S. Horn, *et al.*, "Robot task planning and explanation in open and uncertain worlds," *Artificial Intelligence*, vol. 247, pp. 119–150, 2017.
- [25] G. A. Zachiotis, G. Andrikopoulos, R. Gornez, K. Nakamura, and G. Nikolakopoulos, "A survey on the application trends of home service robotics," in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2018, pp. 1999–2006.

Is Machine Learning Enough to Train Robotic Pets?*

Minqiu Zhou¹, Isobel Voysey^{1,2} and J. Michael Herrmann^{1,2}

Abstract— We discuss the problem of learning in robotic pets asking whether the core machine learning paradigm, namely the optimisation of a bounded error function, is sufficient in this context. In pet robots, it seems that the learning process itself rather than the result of this process is the main criterion for the quality of the interaction. Potential extensions of the optimisation paradigm include emotional, self-organising, and exploratory mechanisms to support desirable learning capabilities of a robotic pet. We also propose a co-design process that develops a personalised interaction experience and mutual learning with active contributions from both robotic pets and their owners.

I. ROBOTIC PETS

Animal companions provide various benefits to their human owners, from longevity and the prevention of coronary heart disease to improved bonding within the family [1]. Pet ownership can be particularly beneficial for children, as the emotional attachments they form with pets can have positive impacts on their socio-emotional development [2]. Owning and caring for a pet can also help to develop children’s understanding of biological concepts, such as inheritance [3].

However, there are many reasons that people do not keep pets, including cost, the responsibility involved, unsuitable housing, and allergies [4], [5]. Robotic pets appear a promising alternative but, despite the early success of Tamagotchi toys [6], most pet-like robots do not reach the level of a companion and are thereby limited to educational support for children [7] or short-term interventions for older adults in care settings [8], [9]. This is despite the potential that older adults living independently see for a robotic pet to enhance their social relationships [10], and it disregards the levels of emotional attachment children show after brief interactions with the robotic dog AIBO [11]. Robotic pets could also be used to teach children (and adults) about various aspects of pet care and appropriate behaviour towards animals. This education prior to the ownership of a live pet, whether in the home or through animal welfare education programmes, would contribute to a reduction of animal suffering [12].

One of the most basic requirements for a robotic or live animal pet is physical companionship. Companion animals are animals kept for the purpose of companionship and comfort [13]. They include a broad range of species, from cats and dogs to reptiles, birds, and fish [14]. Attachments can be formed even with species that have limited capacity for

social interaction [15], and a study comparing users’ feelings of companionship between living, robotic, and virtual pets found that living and robotic pets provided similar levels of companionship, while virtual pets provided markedly less companionship [16], suggesting the mere physical existence of a pet can provide a framework onto which people may project social connections.

For a strong sense of companionship to form with non-human entities, requirements may include an appealing or “cute” appearance [17], a high degree of animacy [18], responsiveness [19], and emotional expression [20], [21], [22]. On a higher behavioural level, it may also be desirable for a robotic pet to provide some level of emotional support, through attunement to the owner’s emotional state [23]. Additionally, the owner may want the pet to display reciprocal attachment, which can be realised by the pet’s behaviour signifying interest and care, such as greeting the owner when they return home [23]. However, each individual will have different motivations and circumstances behind acquiring a pet [24], so the detailed specification of the robotic pet will be highly personal, and individual preferences and level of understanding about animals will need to be taken into account in the design. However, this does not refer a fixed specification of an objective for a pet’s adaptation, as these preferences are developing with the interaction or pet and human, and will not be observable outside the interaction that is to be designed to begin with.

One of the early studies on artificial pets proposes the *uselessness principle* for their design [25]. It suggests that a robotic pet’s primary function is not to provide any service to its owner, thereby necessitating a radically different approach to design compared to the majority of robots. The robotic pet should be driven largely by its own goals and may ignore orders that do not align with its goals. The author argues that this autonomy is a necessary, although not sufficient, feature for the development of an interesting and engaging relationship [25]. In the years since, this idea has been refined in the field of autonomous learning, and in particular in self-motivated [26] and reflexive reinforcement learning [27], skill discovery [28], and others, which we study in the present project.

Recent developments in machine learning enable advanced sensing, planning and action in robotics which we start to appreciate in service robots, but we will argue that an improvement of the information processing capabilities together with increased performance and appearance will not be sufficient to improve the acceptance of robots as pets.

We do not consider the implications of natural language processing here which may be perceived as unnatural in

*This work received partial support from UKRI via CDT RAS Edinburgh.

¹Minqiu Zhou, Isobel Voysey, and J. Michael Herrmann are with the School of Informatics, University of Edinburgh, Edinburgh, EH8 9BT, UK, m.zhou-39@sms.ed.ac.uk, i.a.voysey@sms.ed.ac.uk, michael.herrmann@ed.ac.uk

²Isobel Voysey and J. Michael Herrmann are with the Edinburgh Centre for Robotics, Edinburgh, UK

animal-like robotic pets and has been shown [29] to evoke ambivalent responses in elderly participants, while non-talking robots have a good chance to be perceived as “beautiful” or “compassionate”. We also do not focus on the question of quantifying the performance of robotic pets. There exist subjective measures for people’s feeling of companionship [30] or attachment to pets [31], but in future it will be important to select or develop objective measures to evaluate the performance of machine learning techniques, for example, time spent with the pet.

II. MACHINE LEARNING FOR ROBOTIC PETS

Machine learning methods are yielding increasingly reliable results in tasks with a clearly specified goal, such as path planning or face recognition. This is achieved by the minimisation of a loss function over a data set, or, in the case of reinforcement learning, the maximisation of a reward average. Thus, the optimisation of an objective functions can be seen as the main feature of machine learning algorithms. It has even been claimed that it is possible to design algorithms that realise in this way any form of intelligence [32], if large data sets, complex computational architectures, and sufficiently long learning times are realisable. Yet, we argue that it may not be enough to control a modest pet-like robot, where the eventual result is unimportant.

Although efficient algorithms can achieve few-shot learning or use transfer learning to generalise learned behaviour to new domains, and may produce results comparable to the expectations one might have in an animal pet, it seems that learning success is less important than familiarity, reliability, and some level of creativity. The owner tends to value the time spent with the pet, which is at least partially due to the process of learning itself being more joyful and more important than the final performance of the system.

Error minimisation is, nevertheless, an important component in some of the vital functions of a pet robot as a product, for example, in order to provide basic behaviours (similar to natural traits in a pet animal) and to implement safety regulations. The application of reinforcement learning on a high level, for example to improve user satisfaction, increase engagement duration, and uninterrupted function, is non-trivial, because the exploration of the vast space of potentially useful behaviours requires strategies that need to be designed as well. Therefore, there is a need for active learning mechanisms that can enable a pet robot to find intrinsic motivation to guide reinforcement learning, which can improve smoothness and predictability of movements and supports versatility by skill discovery.

It should also be noted that the unsuccessful execution of a new skill by a pet robot can be appreciated by the owner as it can appear as an intention to learn, to cooperate in this learning process that may be considered as rewarding also to the participating human. This support of the robot’s progress by the human can be considered as a goal of the robot’s learning, such that the cooperative active learning process is characterised by mutual guidance and support. Likewise,

exploratory behaviour will be seen as an attempt to gain information, which can be expected to be fancied in particular if the pet shows a tendency to explore the repertoire of the human. However, to keep user interaction at an acceptable level, it may be suitable to include also emotional dynamics. This has the additional benefit that it can help to supervise the system that is driven by various sources of information in various subsystems. So, a supersystem that monitors the learning progress and the state of the interaction would be useful. This emotional system complements the body of work that has been spent on the recognition and emulation of emotions by the robot. It also provides the drives that control flexibility and adaptability in the learning system that in this way realises a form of self-organisation which, however, deserves further study. It will also support a personalised experience by including sensitivity to a suitable amount of co-operative engagement which needs to be included as a design feature in the pet robot’s control strategy.

III. LEARNING IN PETS AS A CO-DESIGN PROCESS

The development of abilities in a robotic pet should reflect the preferences of the human, but it is not a task to be imposed on them. Instead, the opportunity to continuously adapt the behavioural design of the robot should be given to the human. This and the complementary task of the robot to engage with human, leads to the idea of a co-design process. Co-design is known as a design process where all members contribute in the design process as equal collaborators in a way that fuses optimally the expertise of specialists with the problem-awareness of the users, care-givers, technicians, and other relevant stakeholders. Co-design is common practice in many fields, see e.g. [33], but is considered here a mode of interaction to be realised not between groups of people, but as a dominant mode of human-robot interaction between a pet robot and its owner. This co-design process can be seen as a maturation process that includes various learning processes rather than being simply a learning process itself. Its realisation would include the following points.

Robotic pets are sold with pre-trained sensing capabilities and a few basic behaviours, in other words, the robotic pets arrive in an immature state. After a period of interaction with the environment and the human, the robotic pet gradually grows and develops a unique behavioural organisation, based on the learning experiences they share with the human owners. Research in robotic pets will thus focus on design of co-design which will include the following features.

a) Error tolerance: In contrast to general machine learning that suppresses errors, the maturation process embraces errors which are expected in two ways: In the exploration and self-motivation scenario in reinforcement learning, errors are expected and drive robotic pets. In regard to the use of robotic pets as an entertainment companion, they are allowed to produce slips now and then. Thus, referring to the ‘uselessness’ or ‘cuteness’, some errors, for example in locating owners, could make them appear more alive.

Prior research has shown that forgetfulness might enable a more natural and believable attachment bond between human

and robot companion [34] as opposed to choosing selected “error” behaviours to incorporate, taking thus more holistic view of error tolerance.

b) Personalisation: Individual difference are encouraged in a relationship. With the underlying subjective view, personalisation enriches the human-robot attachment. During the process of maturation, the human-robot interaction becomes more and more specific. It represents a challenge for the behavioural organisation in the robot. In addition to the activation of behaviours by trigger stimuli, any autonomous behaviours need to be grouped according to sequentiality, intended state changes, and context. Although this meta-organisation of behaviour can in principle be learned as well, it may be advisable to provide the robot control architecture with an expressive structure that will become partially inhabited during the maturation process.

c) Active learning: During the maturing period in an individual setting, each robotic pet experiences different interactions and uses any learning successes in the search for new learning data. Thus, in contrast to the assumptions in machine learning, the data sets do not only vary case by case, but are also essentially nonstationary. This leads to complexity and difficulty in the implementation of capturing and defining states or events. Moreover, although robotic pets have access to an enormous number of data during their life, data sets for a specific task are comparably small, which further increases the learning difficulty. However, as errors are not to be avoided, this technical limitation can be experienced as part of the character of the pet.

d) Mutuality: It is not only the robotic pet learning from its owner. Also, owners learn from their robotic pets. This happens while the owners spend time and effort to understand, interpret and control the robotic behaviours. In this way a mutual interaction is formed which is more engaging for the human owners than the mere operation of a machine. Learning in robots and humans works differently in many respects, and it is critical for the robot to be able to access the superior capabilities of humans to adapt. This is possible as demonstrated by the success of computer games.

The implementation of these principles is obviously not straightforward, but can include advances in various fields of machine learning incorporating active learning, imitation learning, transfer learning, reflexive reinforcement learning, and other methods, however, with standards implied by the enjoyable interaction during the learning process, the acceptance of characteristic errors and inconsistencies on both sides, and the development of mutually agreeable repetitive behaviours that would appear as joint rituals.

An evaluation of the quality of a robotic pet as characterised here, would primarily be based on the statistics of the rating of the human user experience. For a more objective account, the duration of daily interactions, and the development of interactions over time can be measured. Beyond this, the complexity of the robotic behaviours can be monitored in pilot cases and information-theoretically analysed. An increase in behavioural complexity with a simultaneous increase of predictability can be seen as in-

dicative of a rich and reliable companionship. It is expressed by the concept of *predictive information* [35] which, in addition to other applications, has been proposed to enable autonomous learning.

IV. DISCUSSION

The project that is described here is still in the making, but it is important to reconsider the principles of the design of a pet robot and to contrast it to the design goals of control architectures of other types of robots. In this way, we can create a niche for a new species of robots that benefits from a symbiosis with its human companions just like the human owners benefit from the robotic companions. Although full functionality of the envisioned pet robot seems a long way ahead, it is still necessary to consider the risks and limitations of this research.

Robotic pets, in particular in applications with elderly persons [8], [9], have been promoted for the purpose of monitoring health, learning progress, and safety, although there are reports of negative side effects of the purposeful sneaky usage of the robot in place of a pet [10]. Similarly, one may object that a fully-functional robotic pet might create a dependency in the interacting human which would not be excusable, *unless* there is a necessary purpose for the presence of the pet in the private domain of the human.

If the robotic pet is used for preparing a (young) person to the responsibilities of owning an animal pet, then any risks of the interaction are contained and will typically be monitored towards the decision of acquiring an animal pet. Likewise, as a temporary companion such as under the conditions of a lockdown or hospital stay the benefits of a robotic pet would outweigh the risks in most cases.

Limitations of the proposed approach consist in the conceptual problem that the robot has no intrinsic need to interact with humans, i.e. the interactive behaviour needs to be explicitly rewarded within the robot’s behavioural module. It may be possible for the pet robot to find out that the interaction does support its intrinsic motivation towards behavioural learning, but then the problem is shifted towards a will to learn. Also, at some point learning progress will saturate due to the limited capabilities of the robot, so that this motivation will diminish.

A related point is the limited function of current hardware. In order for a robot to show impressive behaviour, appropriate hardware needs to be designed, which is beyond the control- and learning-related approach taken here. Needless to mention that here lies much potential for trainable hardware, analogous to muscles, bones, and energy consumption that are known to respond to training in animals.

Although beyond the scope of the present paper, we also need to consider the use of bio-degradable materials, sustainable power sources, as well as minimal-impact behaviours in natural environments.

More generally, the research circumscribed here may be a important component in autonomous robots although probably in combination with trustworthy mechanisms that guarantee a minimal function and limit any risks. It may as

well be useful in setting a scope for the modelling of animal behaviour.

V. CONCLUSION

We discussed the appropriateness of the machine learning techniques for different aspects of robotic pets and concluded that although the machine learning paradigm performs well at sense-and-act levels, it may be too narrow at intentional and cognitive levels. We arrived at the proposal of a co-design learning process for robotic pets, resembling the process of maturation, that cannot simply be learned from rewards, but needs to be accommodated by the design of interaction modes that is an essential part of the design of robotic pets.

REFERENCES

- [1] A. M. Beck and A. H. Katcher, *Between pets and people: The importance of animal companionship*. Purdue University Press, 1996.
- [2] G. F. Melson, *Why the wild things are: Animals in the lives of children*. Harvard University Press, 2005.
- [3] J. M. Williams and L. A. Smith, "Social and experiential influences on the development of inheritance concepts," *International Journal of Behavioral Development*, vol. 30, no. 2, pp. 148–157, 2006.
- [4] K. A. Anderson, L. K. Lord, L. N. Hill, and S. McCune, "Fostering the human-animal bond for older adults: Challenges and opportunities," *Activities, Adaptation & Aging*, vol. 39, no. 1, pp. 32–42, 2015.
- [5] M. C. Gates, J. Walker, S. Zito, and A. Dale, "Cross-sectional survey of pet ownership, veterinary service utilisation, and pet-related expenditures in New Zealand," *New Zealand Veterinary Journal*, vol. 67, no. 6, pp. 306–314, 2019.
- [6] L. Lawton, "Taken by the Tamagotchi: How a toy changed the perspective on mobile technology," *The iJournal: Student Journal of the Faculty of Information*, vol. 2, no. 2, 2017.
- [7] J. D. Decuir, T. Kozuki, V. Matsuda, and J. Piazza, "A friendly face in robotics: Sony's aibo entertainment robot as an educational tool," *Computers in Entertainment (CIE)*, vol. 2, no. 2, pp. 14–14, 2004.
- [8] L. Hung, C. Liu, E. Woldum, A. Au-Yeung, A. Berndt, C. Wallsworth, N. Horne, M. Gregorio, J. Mann, and H. Chaudhury, "The benefits of and barriers to using a social robot PARO in care settings: A scoping review," *BMC Geriatrics*, vol. 19, no. 1, pp. 1–10, 2019.
- [9] S. Petersen, S. Houston, H. Qin, C. Tague, and J. Studley, "The utilization of robotic pets in dementia care," *Journal of Alzheimer's Disease*, vol. 55, no. 2, pp. 569–574, 2017.
- [10] A. Lazar, H. J. Thompson, A. M. Piper, and G. Demiris, "Rethinking the design of robotic pets for older adults," in *Proc. ACM Conf. Designing Interactive Systems*, pp. 1034–1046, 2016.
- [11] A. Weiss, D. Wurhofer, and M. Tscheligi, "'I love this dog' – Children's emotional attachment to the robotic dog aibo," *International Journal of Social Robotics*, vol. 1, no. 3, pp. 243–248, 2009.
- [12] J. M. Williams, M. P. Cardoso, S. Zumaglini, A. L. Finney, S. SPCA, and M. A. Knoll, "'Rabbit rescuers': A school-based animal welfare education intervention for young children," *Anthrozoös*, vol. 35, no. 1, pp. 55–73, 2022.
- [13] A. Steiger, "Pet animals: housing, breeding and welfare," *Animal Welfare*, pp. 111–133, 2006.
- [14] Royal Society for the Prevention of Cruelty to Animals, "RSPCA Policy A01 Responsible companion animal ownership." <https://kb.rspca.org.au/knowledge-base/rspca-policy-a01-responsible-companion-animal-ownership>, 2022. [Online; accessed 11/08/2022].
- [15] J. Langfield and C. James, "Fishy tales: experiences of the occupation of keeping fish as pets," *British Journal of Occupational Therapy*, vol. 72, no. 8, pp. 349–356, 2009.
- [16] J.-L. Rault, "Pets in the digital age: Live, robot, or virtual?," *Frontiers in Veterinary Science*, p. 11, 2015.
- [17] C. Caudwell and C. Lacey, "What do home robots want? The ambivalent power of cuteness in robotic relationships," *Convergence*, vol. 26, no. 4, pp. 956–968, 2020.
- [18] L. Voss, *More than machines?: The attribution of (in) animacy to robot technology*. transcript Verlag, 2021.
- [19] G. E. Birnbaum, M. Mizrahi, G. Hoffman, H. T. Reis, E. J. Finkel, and O. Sass, "Machines as a source of consolation: Robot responsiveness increases human approach behavior and desire for companionship," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 165–172, IEEE, 2016.
- [20] S. M. F. Hosseini, D. Lettinga, E. Vasey, Z. Zheng, M. Jeon, C. H. Park, and A. M. Howard, "Both 'look and feel' matter: Essential factors for robotic companionship," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 150–155, IEEE, 2017.
- [21] N. Martelaro, V. C. Nneji, W. Ju, and P. Hinds, "Tell me more designing hri to encourage more trust, disclosure, and companionship," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 181–188, IEEE, 2016.
- [22] I. Olaronke and I. Rhoda, "A systematic review of emotional intelligence in social robots," 2017. Subm. to International Conference of Applied ICT (ICAICT).
- [23] K. Riddoch, R. Hawkins, and E. S. Cross, "Exploring behaviours perceived as important for human—dog bonding and their translation to a robotic platform," 2021.
- [24] K. E. Holland, "Acquiring a pet dog: A review of factors affecting the decision-making of prospective dog owners," *Animals*, vol. 9, no. 4, p. 124, 2019.
- [25] F. Kaplan, "Free creatures: The role of uselessness in the design of artificial pets," in *Proceedings of the 1st Edutainment Robotics Workshop*, pp. 45–47, Sankt Augustin, 2000.
- [26] N. Chentanez, A. Barto, and S. Singh, "Intrinsically motivated reinforcement learning," *Advances in Neural Information Processing Systems*, vol. 17, p. 1281–1288, 2004.
- [27] B. I. Lyons and J. M. Herrmann, "Reflexive reinforcement learning: Methods for self-referential autonomous learning.," in *IJCCI*, pp. 381–388, 2020.
- [28] B. Eysenbach, A. Gupta, J. Ibarz, and S. Levine, "Diversity is all you need: Learning skills without a reward function," *arXiv preprint arXiv:1802.06070*, 2018.
- [29] S. Coghlan, J. Waycott, A. Lazar, and B. Barbosa Neves, "Dignity, autonomy, and style of company: Dimensions older adults consider for robot companions," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5:CSCW1, no. 104, pp. 1–25, 2021.
- [30] D.-B. Luh, E. C. Li, and Y.-J. Kao, "The development of a companionship scale for artificial pets," *Interacting with Computers*, vol. 27, no. 2, pp. 189–201, 2015.
- [31] T. P. Johnson, T. F. Garrity, and L. Stallones, "Psychometric evaluation of the Lexington attachment to pets scale (LAPS)," *Anthrozoös*, vol. 5, no. 3, pp. 160–175, 1992.
- [32] D. Silver, S. Singh, D. Precup, and R. S. Sutton, "Reward is enough," *Artificial Intelligence*, vol. 299, p. 103535, 2021.
- [33] L. Huybrechts, H. Benesch, and J. Geib, "Institutioning: Participatory design, co-design and the public realm," *CoDesign*, vol. 13, no. 3, pp. 148–159, 2017.
- [34] M. Biswas and J. Murray, "Robotic companionship: how forgetfulness affects long-term human-robot interaction," in *International Conference on Intelligent Robotics and Applications*, pp. 37–48, Springer, 2015.
- [35] W. Bialek, I. Nemenman, and N. Tishby, "Predictability, complexity, and learning," *Neural computation*, vol. 13, no. 11, pp. 2409–2463, 2001.