Enhancing Robotic Manipulation: AR-Powered Data Collection for Learning from Demonstration

Thilina Tharanga Malagalage Don[1], Hamid Khayyam[1], and Ehsan Asadi[1]

[1]School of Engineering, RMIT University

***Abstract*—Integrating robotic manipulators into everyday households faces the significant challenge of allowing them to be taught skills in a natural and humanly understandable way. Although learning-from-demonstration (LFD) shows promise, its reliance on quality data and cumbersome demonstration methods limits its broader application. This paper presents a comparison study on the performance of machine learning models, trained using task demonstration carried out via two traditional methods, two traditional methods augmented with augmented reality (AR), and one augmented reality based method. We compare the performance of these input methods against three ML models and two input data modalities. The results demonstrate the advantage of using AR augmented methods in data collection for LFD and the pure AR method nearly matches the performance of the highest performing AR augmented traditional method while having no drawbacks of the traditional methods.**

***Index Terms*—AR, LFD, ROS, HRI**

I. Introduction

Learning from Demonstration (LfD) seeks to capture and replicate the skills or behaviours shown by humans or machines. Although LfD algorithms excel at mastering short-horizon tasks, they require large amounts of data, and their reliance on expert demonstrations which are considered to be high-quality data poses a challenge. Research suggests that the effectiveness of learned policies is directly related to the quality of the input data, based on the underlying assumption of LFD, that demonstrations represent optimal solutions. However, this assumption often falls short in real-world scenarios. Gathering expert demonstrations is inherently resource intensive and requires significant domain expertise and extensive time investment. This is especially true for tasks involving robotic manipulators with high degrees of freedom (DOFs), where human control of the system presents substantial difficulties even for expert demonstrations due to repeated cognitive and physical demands and constrains.

Current demonstration techniques such as teleoperation[16, 8], passive observation [1], and kinesthetic teaching[5] introduce complexities to the demonstration process, which in turn results in sub-optimal demonstrations[17]. The influence of the human teacher's proficiency on the quality of demonstration data was empirically examined in [10]. As task complexity increases, the demand for precise physical guidance escalates, often exceeding the capabilities of novice human teachers and reducing repeated accuracy of expert teachers.

Augmented Reality (AR) has the potential to reduce the cognitive and physical burden in the demonstration process for LFD [19, 20, 14, 7, 11, 15] as it fundamentally allows

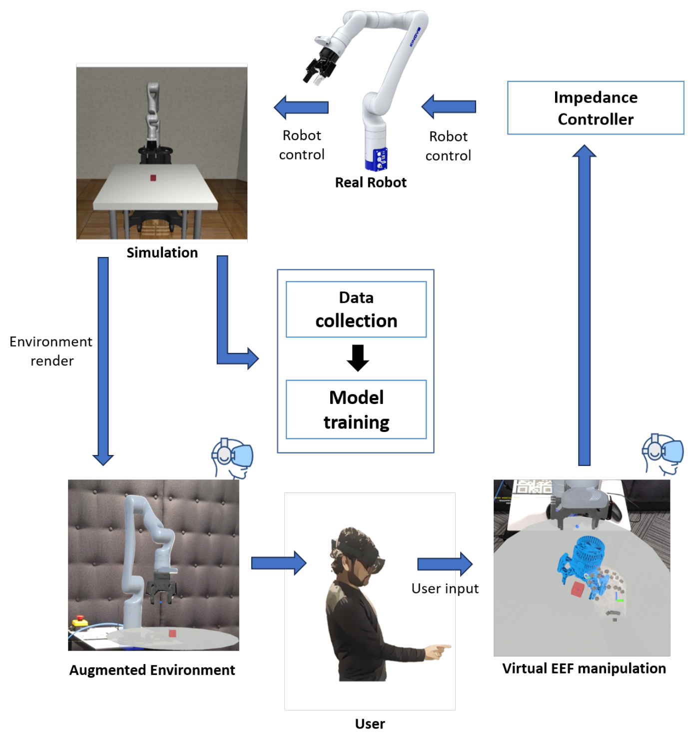

Fig. 1. Architecture: Augmented reality & Real robot blended manipulation teaching environment

to blend the virtual world with the real world seamlessly. AR enables more efficient, accurate, and user-friendly avenues for controlling robots. Specifically, AR has been utilized for reliable visual feedback which reduces the cognitive demand and AR interactive holograms and hologram augmented robots have been used to reduce the physical demand of controlling robots via reducing the forces required and integration of AR supported intuitive high level controllers.

Studies have shown that LFD policy performance can be significantly improved through high quality data which are distinctly characterized by their smother jerk free trajectories and overall and intermediate error free nature. AR based robot arm control methods have shown that they indeed produce trajectories with these qualities thus showing the potential of using AR as an effective tool for generating quality data for LFD policy training demonstrations.

Apart from these, the additional sensing modalities that come with these AR techniques, such as spatial awareness and egocentric sensing, could become the key enablers that solve long-standing issues with LfD techniques.

While existing literature suggests that AR-based demonstration interfaces offer more intuitive, user-friendly, and less cognitively demanding experiences for human demonstrators, there remains a notable gap in empirical research directly comparing LfD performance trained with data collected through various AR enabled input modalities. This comparison is critical, as the effectiveness of LfD systems often hinges on the inherent characteristics of demonstrations, which can vary significantly between these varies input methods.

Moreover data collection methods both traditional and AR powered must be further combined and optimized to address key challenges, including ensuring clarity and ease of demonstration, accommodating higher DOFs without compromising usability, and simplifying the mapping process between the demonstration and the target task. Addressing these challenges is essential to fully leverage the potential of AR and LFD for facilitating the generation of high-quality demonstrations for effective real world applicable AR, LFD and robotic manipulation.

We introduce an AR assisted data collection platform that allows users to collect task demonstrations through 1) conventional methods (kinesthetic, joystick) via visual observation from a display. 2) AR-enabled holographic visual feedback for conventional methods (AR augmented traditional) and 3) AR and impedance control enabled method (AR-live).The system is supported by a popular robotic policy bench mark simulator, Real robotic arm control middle ware, holographic client generating digital twin of the simulation environment and robot in AR and a high speed communication back-end aligning all above subsystems with precision and low latency.

The main contributions of this paper are;

- Demonstration collection system supported by a manipulation benchmark simulator (robosuite) and robot operating system (ROS) based robotic manipulator control backend that streams a holographic digital twin of the simulation tied to a real robotic arm, to a client's AR headset using low-latency communication protocols.Resulting platform is intuitive to use and ties simulation to the real robotic manipulator facilitating realistic benchmark task related evaluations.
- We conduct a study comparing the performance of machine learning models trained using data collected through two traditional, two AR augmented traditional and one novel AR and impedance control facilitated robot arm control method (AR-live), on simulated benchmark object lifting task with task randomization. The introduced novel method performs similarly to augmented traditional methods and significantly outperforms traditional methods.
- We investigates effects of policy selection in relation to these data using three LFD and batch offline RL models and Low dimentional vs image input data.

II. Related Work

End-User Robot Programming Using Mixed Reality evaluated the performance of a Mixed Reality (MR)-based interface for robot programming [3]. The study found that participants were able to program a robot arm to perform pick-and-place tasks more quickly, accurately, and easily using the MR interface than with a traditional 2D interface. The study measured the participants' task completion time, strain, naturalness, and usability, MR interface was found to be significantly better than the 2D interface in terms of all the above criteria. The MR interface was also more usable than the 2D interface. However, the study did not address the generalisation challenges or skill acquisitions combined with MR technologies.

Human-robot interaction for robotic manipulator programming in Mixed Reality (ICRA 2020) present an AR-based robot manipulator trajectory generation system[13, 18, 12, 15]. The system allowed users to define start and end points and use key poses for the robot's trajectory, it also includes path scaling, end effector(EEF) obstacle avoidance, and safety zone visualization[2]. The system features a communication infrastructure built to bridge ROS and unity, where ROS move-it-based manipulator control and unity-based holograms were utilized to define user-friendly, effective manipulator action. The study highlights the potential of combining LFD with the AR control interface introduced as potential future work to address the current skills-learning limitation of the proposed system.

ARC-LfD: Using Augmented Reality for Interactive Long-Term Robot Skill Maintenance via Constrained Learning from Demonstration [9] introduces a kinesthetic teaching-based system augmented by AR that allows users to maintain, update, and adapt learned skills through interaction with the key-frames of learned skills using AR. Users can visualize key frames of the skill, edit them in real-time, and define virtual constraints to guide the robot's relearning process. This, in turn, provides an alternate way to define complex temporal relationships within a demonstration to allow performance enhancement of LDF algorithms. However, the use of kinesthetic teaching for skill acquisition imposes demonstration complexities and requirement for expert demonstrators and the requirement for updating key poses for every environmental change makes system less effective in practical robotic manipulation.

The Benefits of Immersive Demonstrations for Teaching Robots [6] explore the potential of virtual reality (VR) to revolutionize Learning-from-Demonstration (LFD) for robots. The paper demonstrates that VR environments offer an intuitive and efficient way for humans to demonstrate complex tasks, leading to smoother and efficient demonstrations. The study also demonstrates these quality demonstrations results in better performing LFD ML models. The paper compares teleoperation-generated data against VR-generated data over 3 manipulation tasks and concludes the VR data were smoother and shorter, which resulted in learned policies that generated smoother efficient trajectories while requiring fewer data.