Adaptive Robotic Manipulator Simulation for Enhanced Feeding and Drinking Assistance

Soumo Roy
School of Electronics Engineering
Vellore Institute Of Technology
Vellore, India
soumo.roy2021@vitstudent.ac.in

Joel Viju
School of Electronics Engineering
Vellore Institute Of Technology
Vellore, India
joelviju.v2021@vitstudent.ac.in

Budhaditya Bhattacharyya
School of Electronics Engineering
Vellore Institute Of Technology
Vellore, India
budhaditya@vit.ac.in

2024 IEEE 21st India Council International Conference (INDICON) | 979-8-3503-9128-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/INDICON63790.2024.10958301

***Abstract*—In the field of assistive technology, adaptive robotic manipulators offer a promising avenue for improving the standard of living for those who have impairments. This paper details a simulation study of an advanced robotic manipulator developed for feeding and drinking assistance. The simulation features a detailed model of three robotic arm outfitted with a specialized gripper and a spoon-like attachment, showcasing the robot's capacity to adjust to various user requirements and environmental conditions. Key innovations include the robot's real-time adaptation to user-specific feeding angles, precise control of liquid dispensing for drinking, and the creation of intuitive human-robot interaction protocols using reinforcement learning. The system's performance is assessed within a simulated environment that includes a human model interacting with the robot at a dining setup, showing the robot's ability to execute assistive actions with high precision and safety. Preliminary findings suggest that these robotic systems have significant potential to offer reliable and autonomous assistance, thereby enhancing independence and alleviating caregiver workload.**

***Index Terms*—reinforcement learning,assistive robots,openAI gym**

I. Introduction

Robotic technology has revolutionized many areas, with assistive robotics becoming a vital field of research. These robots are designed to help individuals with disabilities, improving their quality of life by assisting with everyday activities. Feeding and drinking assistance are especially important, as they directly affect a person's independence and well-being. This study explores the capabilities of three robotic manipulators—Jaco, Sawyer, and Baxter—in providing such assistance. The research focuses on evaluating the adaptability and efficiency of these robots by comparing three advanced reinforcement learning algorithms: Proximal Policy Optimization (PPO), Soft Actor-Critic (SAC) and Deep Deterministic Policy Gradient (DDPG). Through realistic simulations of human-robot interactions, the study aims to identify the best robot-algorithm combinations, offering valuable insights for the development of more advanced and responsive assistive robotic systems.

II. Literature review

Assistive Gym, an open-source physics simulation framework, models daily tasks and human preferences to optimize robotic assistance. Leveraging reinforcement learning, it has demonstrated enhanced performance in assistive tasks and serves as a valuable tool for advancing research in assistive robotics [1].Improvements were made to a feedback system from sensors for robot assistive pouring tasks by integrating various signals. This optimized feedback setup significantly outperformed standard or random designs, resulting in better task performance and reduced cognitive effort [2].Remote tongue control and semi automation have been applied to assistive robotic arms for individuals with tetraplegia. Findings indicate that semi automation reduces gripping time and command frequency compared to manual control, improving ease of use and independence for users with severe disabilities [3].A hands-free control system combining surface electromyography (sEMG), eye-tracking, and a gyroscope has been developed for operating an assistive robotic arm (ARM) in virtual reality. Results show that individuals with tetraplegia can successfully complete complex tasks with this system, with a clear preference for sEMG controls, enhancing their independence and quality of life [4].Manipulation Primitive Augmented Reinforcement Learning (MAPLE) advances reinforcement learning through the integration of behaviour primitives, leading to enhanced performance in simulated manipulation tasks and effective implementation in real-world scenarios [5].An effective framework is offered for using Deep Reinforcement Learning (DRL) in robotic arm control, emphasizing the integration of simulations using the Robosuite platform and essential tools, along with possible industrial uses like bin picking operations for warehouses [6]. The development and optimization of FOODIEBOT, a food delivery robot, involved several algorithms (BAS, PSO, POA, EO) to enhance its performance across different paths. The BAS method was found to be superior in accuracy and execution time, while other algorithms varied in speed and path efficiency, confirming the real-world applicability of the simulation results [7].A DQN-based reinforcement learning approach is proposed for solving 7-DOF manipulator IK, offering a stable, efficient alternative for generating joint-space trajectories and adaptable across robotic systems [8].A deep reinforcement learning (DRL) method improves trajectory planning for manipulators in dynamic environments, using a dynamic action selection strategy and a combined reward function to boost convergence up to 3-5 times in Deep Deterministic Policy Gradient (DDPG), Twin Delayed Deep Deterministic Policy Gradient (TD3), and Soft Actor-Critic (SAC) algorithms [9].A novel framework and metrics are introduced to select and assess Deep Reinforcement Learning (DRL) models for dynamic pricing [10].A robot-assisted feeding system for mobility impaired individuals is demonstrated, ensuring safety, portability, and user control, with a custom web-app for control and real-time bite transfer [11].

A. Research context

Existing reinforcement learning systems need to be generalized to work across different users and various types of food, ensuring consistent performance.The balance between safety, portability, and performance remains a challenge. There is a need for robust systems that can ensure user safety without compromising on portability and operational efficiency. Improvement in real-time error handling mechanisms is crucial to address unexpected issues during feeding tasks and drinking task, ensuring seamless operation.

III. Methodology

The simulations are performed in the Assistive Gym environment, which offers realistic physical interactions and detailed human-robot collaboration models. The environment includes three robotic arms—Jaco, Baxter, and Sawyer robots are chosen for simulations due to their adaptable design and precision in assistive tasks, which enable high reward functions. Virtual representations of individuals with physical disabilities are used to simulate the human aspect, while the dining setup mimics real-world conditions with a table, chair, and feeding apparatus.

A. Robots

1) **Jaco :** The Jaco robot, designed for assistive applications, features a lightweight yet robust construction, emphasizing precision and adaptability. Its kinematic design includes a 6-degree-of-freedom (DOF) arm, which allows for a broad range of motion and dexterity, critical for tasks requiring fine manipulation.The gripper is engineered with customizable fingers and a specialized spoon-like attachment, offering nuanced control suited for feeding tasks

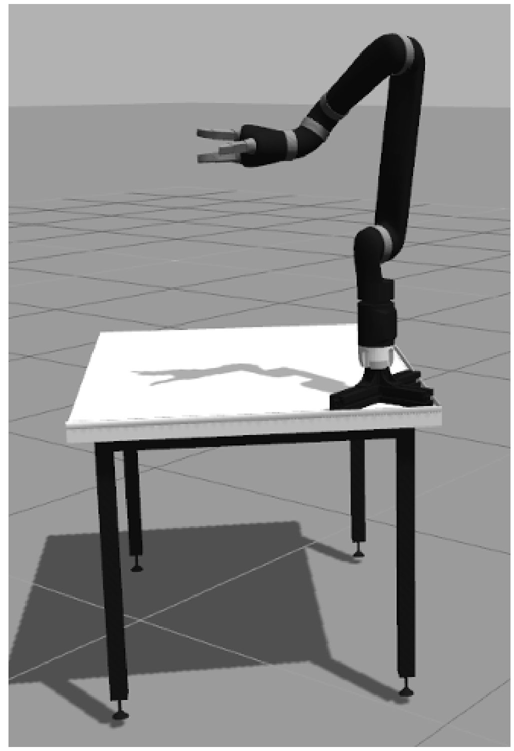

Fig. 1. Jaco robot

2) **Baxter :** The Baxter robot, designed for both industrial and research environments, features a dual-arm system with 7 degrees of freedom (DOF) per arm. Each arm is equipped with advanced joint torque sensors and high-resolution encoders for detailed feedback and accurate movement. It comes with a custom gripper and a spoon-like tool, both incorporating force-sensitive resistors and motorized adjustments to handle delicate feeding tasks.It is mounted on a mobile base with omni-directional wheels.

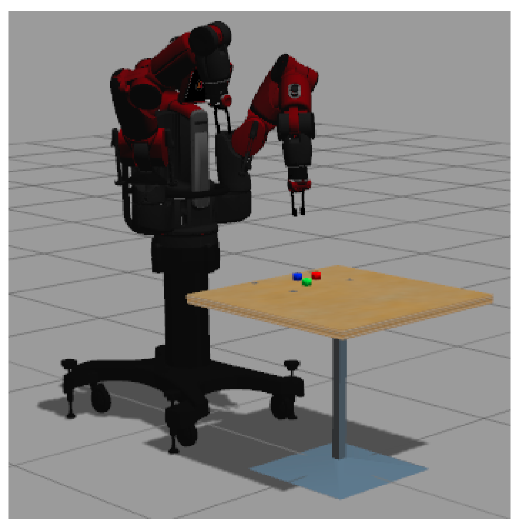

Fig. 2. Baxter robot

3) **Sawyer :** The Sawyer robot, celebrated for its precision and versatility, is a single-arm manipulator designed for complex tasks. It has a 7-degree-of-freedom (DOF) arm equipped with advanced force-torque sensors and high-resolution encoders for detailed control and movement feedback. Sawyer's gripper and a variety of tools allow it to adjust to different tasks, including feeding and drinking assistance.

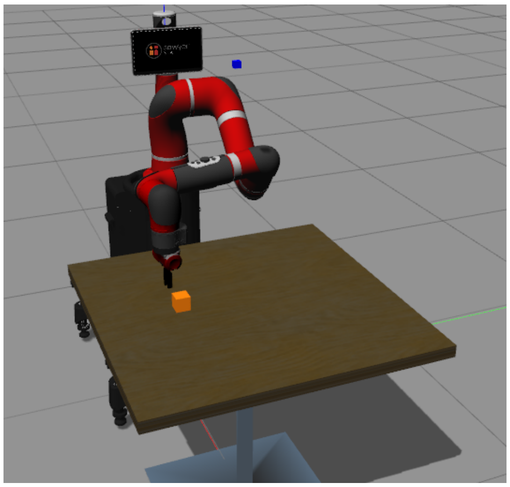

Fig. 3. Sawyer robot

B. Algorithm

1) **PPO :** The Proximal Policy Optimization(PPO) uses a neural network policy and a value network to optimize the Proximal Policy Optimization algorithm, ensuring stable learning and accurate estimation of the value function.

Algorithm 1 Proximal Policy Optimization

for $i \in \{1, \cdots, N\}$ **do**
  Set the policy π_θ for T timesteps, collecting $\{s_t, a_t, r_t\}$
  Predicting the advantages $\hat{A}_t = \sum_{t'>t} \gamma^{t'-t} r_{t'} - V_\phi(s_t)$
  $\pi_{\text{old}} \leftarrow \pi_\theta$
  for $j \in \{1, \cdots, M\}$ **do**
    $J_{PPO}(\theta) = \sum_{t=1}^{T} \frac{\pi_\theta(a_t|s_t)}{\pi_{old}(a_t|s_t)} \hat{A}_t - \lambda \text{KL}[\pi_{old}|\pi_\theta]$
    Update θ by a gradient method w.r.t. $J_{PPO}(\theta)$
  end for
  for $j \in \{1, \cdots, B\}$ **do**
    $L_{BL}(\phi) = -\sum_{t=1}^{T} (\sum_{t'>t} \gamma^{t'-t} r_{t'} - V_\phi(s_t))^2$
    Update ϕ by a gradient method w.r.t. $L_{BL}(\phi)$
  end for
  if $\text{KL}[\pi_{old}|\pi_\theta] > \beta_{high} \text{KL}_{target}$ **then**
    $\lambda \leftarrow \alpha\lambda$
  else if $\text{KL}[\pi_{old}|\pi_\theta] < \beta_{low} \text{KL}_{target}$ **then**
    $\lambda \leftarrow \lambda/\alpha$
  end if
end for

2) **SAC :** The Soft Actor-Critic algorithm(SAC) optimizes policy networks, promoting exploration, by mapping observations to probability distributions over actions, and automatically adjusting temperature parameters to maintain balance.

Algorithm 2 Soft Actor-Critic

Set the parameter vectors to ψ, $\bar{\psi}$, θ, ϕ.
for each loop **do**
  for each environment step **do**
    $a_t \sim \pi_\phi(\mathrm{a}_t|\mathrm{s}_t)$
    $s_{t+1} \sim p(s_{t+1}|s_t, a_t)$
    $\mathcal{D} \leftarrow \mathcal{D} \cup \{(s_t, a_t, (s_t, a_t), s_{t+1})\}$
  end for
  for each gradient step **do**
    $\psi \leftarrow \psi - \lambda_V \hat{\nabla}_\psi J_V(\psi)$
    $\theta_i \leftarrow \theta_i - \lambda_Q \hat{\nabla}_\theta J_Q(\theta_i)$ for $i \in \{1, 2\}$
    $\phi \leftarrow \phi - \lambda_\pi \nabla_\phi J_\pi(\phi)$
    $\bar{\psi} \leftarrow \tau\phi + (1-\tau)\bar{\psi}$
  end for
end for

3) **DDPG :** Deep Deterministic Policy Gradient (DDPG) is a reinforcement learning method for continuous actions, involving actors and critics, storing past experiences, updating networks, and introducing noise for exploration.

Algorithm 3 Deep Deterministic Policy Gradient

Set the parameter vectors to $Q(s, a|\theta^Q)$ and actor $\mu(s|\theta^\mu)$ with weights θ^Q and θ^μ.
Set the target network Q' and μ' with weights $\theta^{Q'} \leftarrow \theta^Q$, $\theta^{\mu'} \leftarrow \theta^\mu$
Set the replay buffer R
for episode = 1, M **do**
  Set a random process $\mathcal{N}$ for action investigation
  Get initial observation state s_1
  for t = 1, T **do**
    Select action $a_t = \mu(s_t|\theta^\mu) + \mathcal{N}_t$ according to the current policy and investigation noise
    Set the action a_t and observe reward r_t and see the new state s_{t+1}
    Store transition (s_t, a_t, r_t, s_{t+1}) in R
    Sample a random small batch of N transitions (s_i, a_i, r_i, s_{i+1}) from R
    Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$
    Update critic by reducing the loss: $L = \frac{1}{N}\sum_i (y_i - Q(s_i, a_i|\theta^Q))^2$
    Update the actor policy with the sampled policy gradient:

$$\nabla_{\theta^\mu} J \approx \frac{1}{N} \sum_i \nabla_a Q(s, a|\theta^Q)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s|\theta^\mu)|_{s_i}$$

    Update the target networks:

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$$

$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$$

  end for
end for

C. Human posture study

We observed the human to be seated on a wheelchair as we wanted to provide feeding and drinking assistance to patients with Parkinson disease as under our observation most of these patients tend to be seated on a wheelchair in hospitals when they are being treated by a nurse.

IV. Results

A. Simulation output

In the Fig.4. we can observe the simulation output for Jaco robot providing feeding assistance to the patient.

Fig. 4. Jaco feeding assistance simulation

In the Fig.5. we can observe the simulation output for Baxter robot providing feeding assistance to the patient.

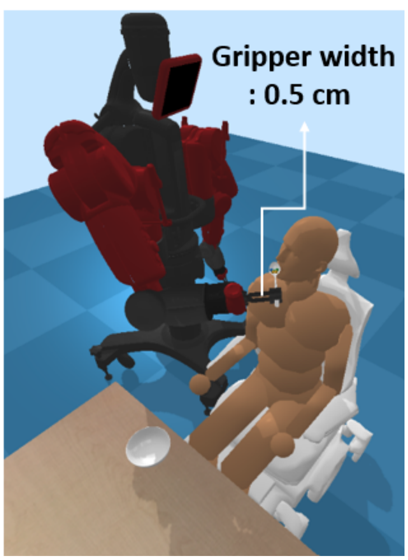

Fig. 5. Baxter feeding assistance simulation

In the Fig.6. we can observe the simulation output for Sawyer robot providing feeding assistance to the patient.

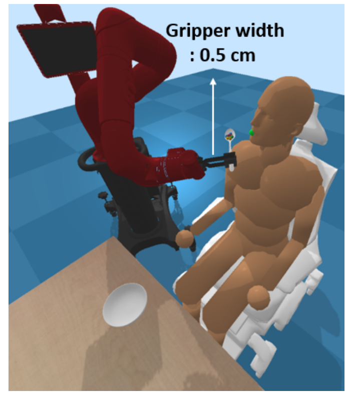

Fig. 6. Sawyer feeding assistance simulation

In the Fig.7. we can observe the simulation output for Jaco robot providing drinking assistance to the patient.

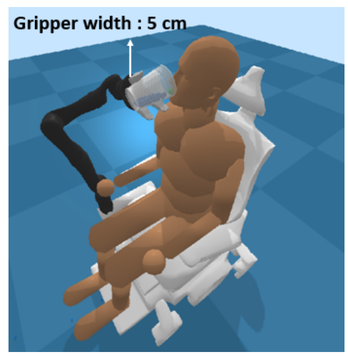

Fig. 7. Jaco drinking assurance simulation

In the Fig.8. we can observe the simulation output for Baxter robot providing drinking assistance to the patient.

Fig. 8. Baxter drinking assistance simulation

In the Fig.9. we can observe the simulation output for Sawyer robot providing drinking assistance to the patient.

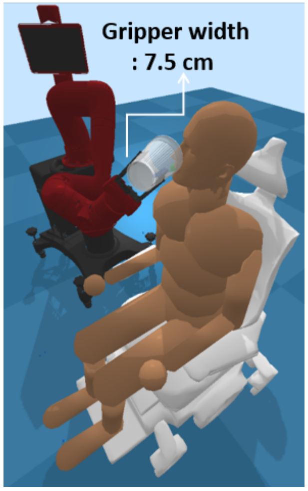

Fig. 9. Sawyer drinking assistance simulation

B. Tables

For 10,000,000 time steps, or 50,000 simulation rollouts (trials), we use 36 concurrent simulation actors to train each policy. In a simulated rollout, a policy can perform a new action at every 200 time steps (20 seconds of time for simulation at 10 time steps per second). Every 7,200 time steps, or when each actor finishes a single simulation trial, we update the policy every 10 epochs.

TABLE I
AVERAGE REWARDS AND TASK SUCCESS FOR 100 TRIALS FOR FEEDING ASSISTANCE TASK

Algorithm	*Jaco*	*Baxter*	*Sawyer*	*Success*
PPO	83.6	108.1	95.4	87%
SAC	105.1	107.3	110.5	88%
DDPG	130	98.7	107.6	83%

TABLE II
AVERAGE REWARDS AND TASK SUCCESS FOR 100 TRIALS FOR DRINKING ASSISTANCE TASK

Algorithm	*Jaco*	*Baxter*	*Sawyer*	*Success*
PPO	85.7	263.3	436.0	72%
SAC	402.6	466.8	464.0	73%
DDPG	108	371.88	441.2	78%

The reward function reflects how well different algorithms (PPO, SAC, DDPG) perform in feeding and drinking assistance tasks with various robots (Jaco, Baxter, Sawyer).The results show that the success rate and rewards vary depending on the task and the robot used, with an average success rate of 86% for feeding and 74.33% for drinking, highlighting the importance of selecting the right algorithm for specific tasks.

C. Graph

Fig 10,11,12,13,14,15 shows the graphs for validation of task success and food rewards for each robot in different feeding and drinking scenarios.

Fig. 10. Jaco feeding assistance graph

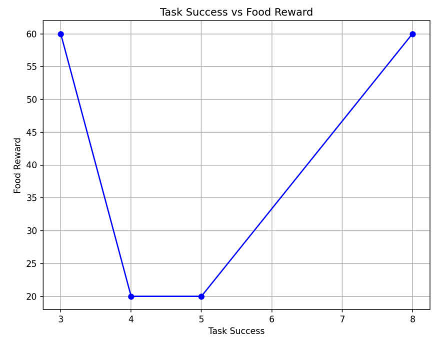

Fig. 11. Baxter feeding assistance graph

Fig. 12. Sawyer feeding assistance graph

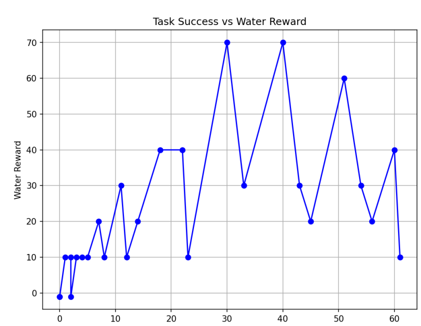

Fig. 13. Jaco drinking assistance graph

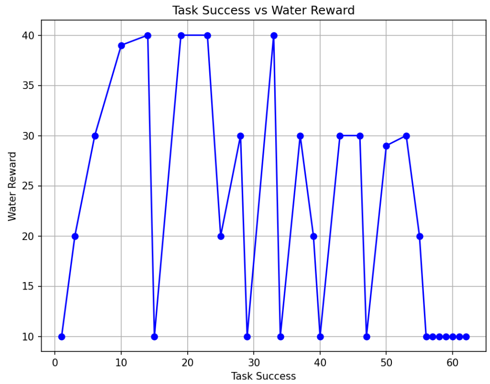

Fig. 14. Baxter drinking assistance graph

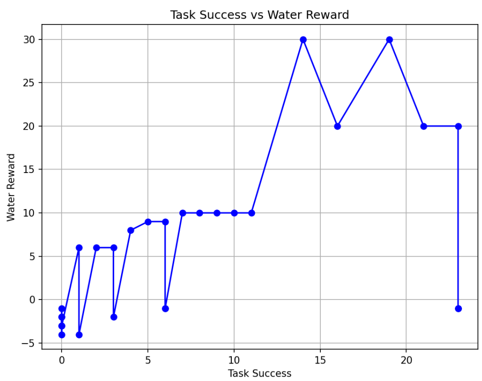

Fig. 15. Sawyer drinking assistance graph

The connection between task accomplishment and reward is intricate and diverse, varying greatly across different situations. These complex patterns likely stem from a mix of different reward systems and experimental setups.

CONCLUSION

This study has systematically evaluated the performance of three robotic manipulators—Jaco, Sawyer, Baxter—in the context of assistive feeding and drinking tasks. By implementing and comparing the Proximal Policy Optimization (PPO) ,Soft Actor-Critic (SAC) algorithm and Deep Deterministic Policy Gradient(DDPG), we have identified the relative strengths and weaknesses of each robot-algorithm pair. Our findings indicate that while some robots excel in precision and safety, others demonstrate faster adaptation to task variations. The comparative analysis reveals that no single algorithm consistently outperforms the other across all robots, stating that the algorithm chosen should be customized for the particular robot and task requirements. Overall, this study contributes to the ongoing efforts to improve the capabilities of assistive robots, potentially enhancing the standard of living for those who have physical disabilities.

We would like to transition from simulation to real-world environments to validate the findings under more complex and unpredictable conditions using robots built using custom hardware which would be lightweight in nature instead of standard robots like Jaco,Baxter and Sawyer. This will involve addressing challenges such as sensor noise, actuation delays, and physical interactions with humans. Introducing real-time monitoring or fail-safe mechanisms to prevent accidents would be an important area to consider in future work.

ACKNOWLEDGMENT

We would like to thank our professor Dr. Budhaditya Bhattacharya for his support, guidance and valuable time. This work was financialy supported by Vellore Institute of Technology (VIT), Vellore under the Faculty Seed Grant (RGEMS) (Sanction Order No.: SG20220001).

REFERENCES

[1] Z. Erickson, V. Gangaram, A. Kapusta, C. K. Liu, and C. C. Kemp, “Assistive Gym: a physics simulation framework for assistive robotics,” 2020 IEEE International Conference on Robotics and Automation (ICRA), May 2020, doi: 10.1109/icra40945.2020.9197411.

[2] M. S. Marambe, B. S. Duerstock, and J. P. Wachs, “Optimization approach for multisensory feedback in Robot-Assisted pouring Task,” Actuators, vol. 13, no. 4, p. 152, Apr. 2024, doi: 10.3390/act13040152.

[3] Á. A. Pálsdóttir, R. L. Kæseler, S. H. Bengtson, T. B. Moeslund, and L. N. S. A. Struijk, “Remote tongue-based control of a wheelchair mounted assistive robotic manipulator and the effect of adaptive semi-automation,” IEEE Transactions on Biomedical Engineering, pp. 1–10, Jan. 2024, doi: 10.1109/tbme.2024.3435837.

[4] T. C. Hansen, T. N. Tully, V. J. Mathews, and D. J. Warren, “A multimodal Assistive-Robotic-Arm control System to increase independence after tetraplegia,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 32, pp. 2124–2133, Jan. 2024, doi: 10.1109/tnsre.2024.3408833.

[5] S. Nasiriany, H. Liu, and Y. Zhu, “Augmenting Reinforcement Learning with Behavior Primitives for Diverse Manipulation Tasks,” 2022 International Conference on Robotics and Automation (ICRA), May 2022, doi: 10.1109/icra46639.2022.9812140.

[6] C. Calderón-Cordova, R. Sarango, D. Castillo, and V. Lakshminarayanan, “A deep reinforcement learning framework for control of robotic manipulators in simulated environments,” IEEE Access, vol. 12, pp. 103133–103161, Jan. 2024, doi: 10.1109/access.2024.3432741.

[7] A. J. Moshayedi, A. S. Roy, L. Liao, A. S. Khan, A. Kolahdooz, and A. Eftekhari, “Design and Development of Foodiebot Robot: from Simulation to Design,” IEEE Access, p. 1, Jan. 2024, doi: 10.1109/access.2024.3355278.

[8] A. Malik, Y. Lischuk, T. Henderson, and R. Prazenica, “A Deep Reinforcement-Learning approach for inverse Kinematics solution of a high degree of freedom robotic manipulator,” Robotics, vol. 11, no. 2, p. 44, Apr. 2022, doi: 10.3390/robotics11020044.

[9] L. Zheng, Y. Wang, R. Yang, S. Wu, R. Guo, and E. Dong, “An efficiently convergent Deep Reinforcement Learning-Based Trajectory planning method for manipulators in dynamic environments,” Journal of Intelligent & Robotic Systems, vol. 107, no. 4, Mar. 2023, doi: 10.1007/s10846-023-01822-5.

[10] Y. Liu, K. L. Man, G. Li, T. R. Payne, and Y. Yue, “Evaluating and Selecting Deep Reinforcement Learning Models for OptimalDynamic Pricing: A Systematic Comparison of PPO, DDPG, and SAC,” CCEAI ’24: Proceedings of the 2024 8th International Conference on Control Engineering and Artificial Intelligence, Jan. 2024, doi: 10.1145/3640824.3640871.

[11] E. K. Gordon et al., “An Adaptable, Safe, and Portable Robot-Assisted Feeding System,” HRI ’24: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, Mar. 2024, doi: 10.1145/3610978.3641085.