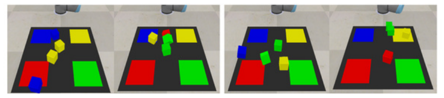
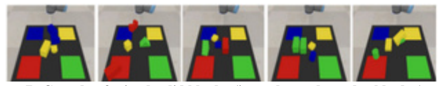
Designing an Intelligent Self-Learning Robot for Adaptive Object Sorting

Research Problem

- The fifth industrial revolution emphasizes human-robot collaboration, highlighting the need for adaptive robotic systems in education [1] and industrial application domains
- Training an agent to generate cooperative joint learning policies for executing object sorting in a cluttered scene is very complex
- How can a customized PQCN agent enable sorting of cubic and mixed objects in dynamic environments?



A. Sample of cubic solid blocks



B. Sample of mixed solid blocks (irregular and regular blocks)

Figure 1: Examples of observation scene containing either cubic or irregular solid objects

Scientific Background

- Reinforcement learning (RL) in robotic object sorting has progressed through deep reinforcement learning (DRL) algorithms, such as DDPG and Deep Q-Networks [2], applied in simulation platforms like VREP (CoppelliaSim), OpenAI Gym, MuJoCo, and Gazebo
- Adapting to dynamic real-world environments remains an ongoing research challenge. Google's RL agent, Predictive Information-QT-Opt (PI-QT-Opt) [3], achieved an 84% accuracy rate when deployed in a robotic waste-sorting task

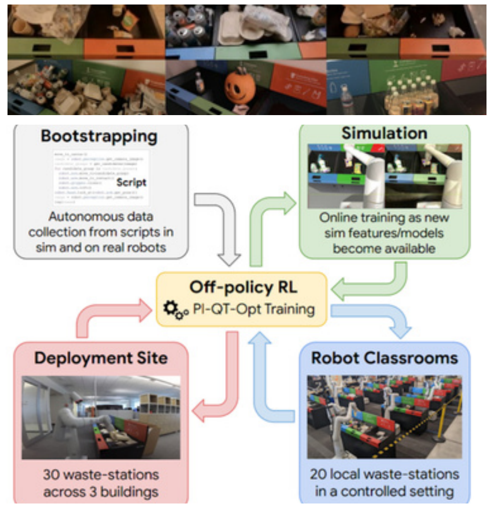


Figure 2: Practical application of multiple robots using an off-policy RL used for waste sorting [3]

- Pixel-wise Critic Networks with Q-values (PQCN) are used for visual-based sorting, despite being computationally demanding [4]
- The research gaps include challenges in scaling from simulated to real-world environments and the high complexity of resource-intensive models, such as DenseNet121 and MobileNetV3-Large. These models were primarily utilized for sorting mixed-object blocks but did not address the development of models specifically designed for sorting cubic blocks alone【4】.

Objectives

- Develop a PQCN using MobileNetV3-Small to reduce critic network complexity by 40% and optimize sorting policies
- Explore two PQCN models for sorting cubic and irregular solid blocks in dynamic environments
- Identify the model with better generalization for complex object sorting during testing

Hypothesis

Explore two PQCN models for sorting regular and irregular blocks in dynamic environments

- Identify the model with better generalization for handling complex object variations

Variables

- Independent Variables: State observations and two DRL models for sorting cubes and irregular blocks in four colors (red, blue, yellow, green)
- Dependent Variables: Performance metrics: AE, GSR, PSR, and SSR
- Controlled Variables: dynamic environmental conditions (sparse, cluttered, and stacked objects) during training and testing

Methodology

System Pipeline

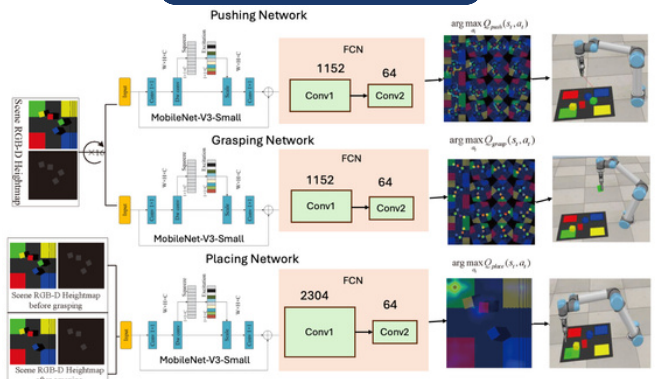


Figure 3: The proposed deep reinforcement learning system for performing object sorting while executing three main tasks: pushing, grasping, and placing

Procedures

- Design: A lightweight Pixel-wise Q-valued critic network with MobileNetV3-Small and a condensed FCN, trained using stochastic gradient descent optimization scheme
- Training: The agent was trained for 25,000 time step on dynamic environments with irregular and cubic blocks for object sorting
- Testing: Models were tested on cubic blocks with different colors, using metrics: AE, GSR, PSR, and SSR
- Comparison: The models were compared for generalization based on standard performance metrics

Graph & Data

Graphs Results

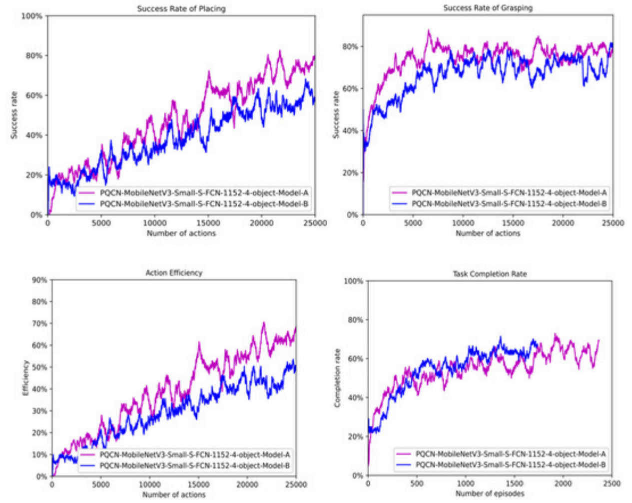


Figure 4. Training phase performance index measured from PQCN-Mobilenet-V3-Small-FCN-1152optimized with stochastic gradient descent with momentum after training for 25,000 action steps.

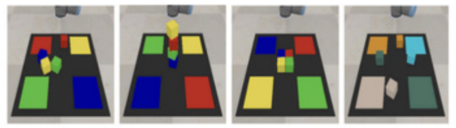


Figure 5: Testing environment: generated from fixed observation scenes that shows 4 object blocks have unique sets of challenges: sparse, clutter, occluded, and variation of object colors

Tabular Results

Table 1: PQCN-MobileNetV3-Small-FCN-1152-Model-A evaluated on fixed testing environments

Performance Metrics	Test 1	Test 2	Test 3	Test 4	Mean
Action Efficiency	79.25	38.11	35.44	27.53	45.08 ± 20.11
Grasp Success Rate	78.66	64.21	39.84	73.67	63.35 ± 14.94
Place Success Rate	95.64	56.57	69.14	32.51	63.46 ± 22.77
Sort Success Rate	91.12	45.16	75.00	65.00	69.07 ± 16.66

Table 2: PQCN-MobileNetV3-Small-FCN-1152-Model-B evaluated on fixed testing environments

Performance Metrics	Test 1	Test 2	Test 3	Test 4	Mean
Action Efficiency	70.73	63.18	44.23	35.46	53.40 ± 14.16
Grasp Success Rate	76.19	60.61	46.49	60.81	61.03 ± 10.51
Place Success Rate	83.14	85.27	72.52	51.80	73.18 ± 13.26
Sort Success Rate	95.00	80.83	80.00	65.83	80.41 ± 10.32

Results

- Training: PQCN-MobileNetV3-Small-FCN-1152 models were trained for 25,000 timesteps on cubic (Model A) and irregular blocks (Model B).
- Performance: Model A excelled in GSR, PSR, and AE but had lower sample efficiency than Model B.
- Test Results: Model B outperformed Model A in 3 out of 4 metrics, with higher SSR, PSR, and AE.
- Adaptability: Model B showed better generalization in complex environments (cluttered, varied object colors).
- Recommendation: Model B is preferred for real-robot deployment due to superior sample efficiency and adaptability.

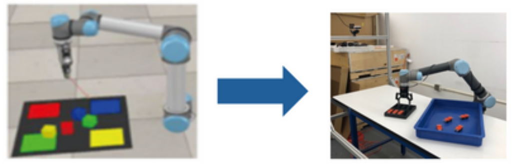
Conclusion

- Proposed Efficient PQCN Design: Developed a 40% simplified PQCN architecture using MobileNetV3-Small and a compact fully convolutional network to enhance object sorting efficiency
- Task Optimization: Implemented multi-step learning for pushing, grasping, and placing, with separate models for sorting cubic and mixed irregular blocks
- Generalization Insight: Model B demonstrated superior handling of complex, unseen environments during testing, making it the preferred choice for deployment

Applications

Figure 5: AI-powered industrial robots performing quality assurance, packaging, and shipping to enhance efficiency and reduce waste (products from Tulips Technologies).

Future Work



- Deployment: Integrate the agent with a real-time UR5 robot
- Expansion: Extend the self-learning robot to industrial applications.

References

[1] Yang, Q. F., Lian, L. W., & Zhao, J. H. (2023). Developing a gamified artificial intelligence educational robot to promote learning effectiveness and behavior in laboratory safety courses for undergraduate students. International journal of educational technology in higher education, 20(1), 18.

[2] Bao, J., Zhang, G., Peng, Y., Shao, Z., & Song, A. (2022). Learn multi-step object sorting tasks through deep reinforcement learning. Robotica, 40(11), 3878-3894.

[3] Herzog, A., Rao, K., Hausman, K., Lu, Y., Wohlhart, P., Yan, M., ... & Levine, S. (2023). Deep rl at scale: Sorting waste in office buildings with a fleet of mobile manipulators. *arXiv preprint arXiv:2305.03270*.

[4] Okafor, E., Oyedeji, M., & Alfarraj, M. (2024). Deep reinforcement learning with light-weight vision model for sequential robotic object sorting. Journal of King Saud University-Computer and Information Sciences, 36(1), 101896.