ICRA2026 Late Breaking Results Poster presented at
2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Distributed AI for Robotics

Satyabhama Singh, Lars Wulfert, Hendrik Wöhrle
Embedded Software and AI
Fraunhofer Institute for Microelectronic Circuits and Systems
Duisburg, Germany
{satyabhama.singh, lars.wulfert, hendrik.woehrle}@ims.fraunhofer.de

***Abstract*—Robot learning primarily relies on centralized training. While it provides the infrastructure, centralization limits parallel and collaborative learning among robots and place significant computational load on the central server, indicating the need for federated learning (FL) in context of multi-robot training. However, robots trained in a federated setup are subjected to non-independent and identically distributed data (non-IID), resulting in degraded model performance. This extended abstract presents the current state of research aimed at improving robot learning under non-IID conditions in FL. In this regard, this work provides an initial comparative analysis of robot learning methods in centralized and federated training setups, with an emphasis on the impact of non-IID data on learning behaviour in a simulation environment. The results highlight the differences in learning stability across algorithms and present the influence of non-IID goal distributions on performance.**

Keywords—federated learning, deep reinforcement learning, non-IID, robot manipulation

I. INTRODUCTION

Over the years, robot learning has evolved from rule-based approaches to more adaptive artificial intelligence (AI)-based methods for complex environments. However, AI-based algorithms are often trained in a centralized manner, where a single computing server is responsible for orchestrating the training process and storing data. While this setup is effective, robot learning algorithms require large datasets and are compute-intensive, adding significant load to the server. Moreover, from a multi-robot learning perspective, this setup lacks parallel and collaborative learning among robots. These limitations indicate the need for distributed training.

In this context, FL enables distributed training, wherein multiple robots learn locally over their own dataset while sharing knowledge (i.e., model parameters rather than raw data) through a central server, thereby supporting parallel and collaborative training of multiple robots. However, FL has challenges such as non-independent and identically distributed data (non-IID), communication constraints, and model convergence.

In this regard, our research goal is to study robot learning under non-IID, guided through the following research questions:

- **RQ1: Centralized vs. federated learning:** How does the trade-off between performance and training efficiency vary across federated and centralized training of robot behaviour models for manipulation tasks?
- **RQ2: Robot learning under non-IID:** How can data-based strategies improve robot learning under non-IID conditions, and what impact do they have on model generalization?

The current scope is to present research at the intersection of robotics and FL, define non-IID in federated robot learning, and present initial results of robot learning trends in sparse reward environments under centralized and federated settings, with an emphasis on observing the impact of non-IID on robot learning. These initial experiments address the first research question through preliminary insights, and the remaining questions are considered as part of our future work.

II. STATE-OF-THE-ART

The current state of research at the intersection of FL and robotics have been approached from multiple perspectives, including performance and communication challenges, data heterogeneity, and emerging benchmarks. Approaches such as MTF-Grasp [1] and FedAgent [2] address robot learning through multi-tier and device selection strategies, respectively, while methods like FORLA [3] focus on handling non-IID through improved representation learning in vision-based tasks. FLAME [4] presents the first benchmark in FL for robotic manipulation through imitation learning. In the area of generative AI, FedVLA [5] integrates FL with Vision-Language-Action models. Ma et al. [6] presents survey focused on non-IID solutions and classify these methods into data-based, model-based, algorithm-based, and framework-based strategies. Prior works have also explored efficient resource utilization through cloud-based robotics [7] and decentralization, including SDRL [8].

III. NON-IID IN FEDERATED ROBOT LEARNING

Robot learning degrades under non-IID across multiple clients. In federated robot learning, non-IID stems from robots being subject to variations in the environment and training conditions. When trained under heterogeneity, the aggregated knowledge at the server affects two main aspects of robot learning, i.e., the observations and the actions they take for a given task.

Non-IID is formally represented by statistical heterogeneity. In this work, it is formulated through differences in joint distribution across observations, actions, and rewards. A random selection of client devices i and j, is expressed as $P_i(x, y)$, where x and y present actions and observations, respectively. The table represents the resulting skewness under variations in x and y, following are their conditional probability representations:

- **Observation skew *(x | y)*:** $p_i(x|y) \neq p_j(x|y)$
- **Action skew *(y | x)*:** $p_i(y|x) \neq p_j(y|x)$

- **Experience heterogeneity *(x, y)*:** $p_i(x, y) \neq p_j(x, y)$
- **Independent and identically distributed (IID) *(x, y)*:** $p_i(x, y) = p_j(x, y)$
- **Rewards skew *(x, y, r)*:** $p_i(x, y, r) \neq p_j(x, y, r)$

IV. Initial results and discussion

This section presents initial findings from training and evaluation of DRL methods in a robot simulation environment, analyzed through their reward convergence and success rates in centralized and federated training setups under IID and non-IID scenarios, and their influence on robot learning.

A. Task definition

At this stage of research, a simple manipulation task is considered where the gripper is required to reach a target goal position. For this purpose, the *FetchReach-v3* [9] environment is used. The observations consist of the robot states, desired goal, and achieved goal, and the model output is a continuous action representing Cartesian movement of the gripper in 3D space.

B. Model selection for experimentation

Off-policy actor-critic methods, i.e., Soft Actor Critic (SAC) [10] and Deep Deterministic Policy Gradient (DDPG) [11], are considered. In sparse reward robotic environments, Hindsight Experience Replay (HER) [12] enables failures to contribute to policy learning. This work integrates HER with both SAC and DDPG and this approach is extended to a federated setting.

C. Implementation setup

Federated training is set up by simulating a server with three client devices. SAC+HER and DDPG+HER are evaluated in centralized and federated settings through reward convergence and success rate over 50 episodes. Both policies are distributed across clients while HER is set local to each client. Non-IID is introduced via heterogeneous goal distribution across clients. Models are trained for 1M timesteps in the centralized setting and 50K per client over 10 communication rounds in federated setup, and global aggregation is performed using FedAvg [13].

D. Initial Results

1) Centralized training: SAC + HER achieved stable and consistent training with a steady improvement in reward and success rate. In comparison, DDPG + HER also indicated an improvement in learning; however, it was highly sensitive to exploration, with higher variability and lower success rates across multiple runs.

2) Federated training with IID: Both SAC + HER and DDPG + HER showed consistent improvement in reward and success rate over timesteps. While a direct comparison to the centralized setup is limited due to different training timesteps in this initial study, similar learning trends were observed in the IID setting. SAC + HER demonstrated relatively stable and consistently higher performance across episodes, whereas DDPG + HER indicated comparatively higher variability.

3) Federated training with non-IID: Introducing goal distribution heterogeneity resulted in a degraded performance for both methods compared to the IID setting, as reflected by a drop in their rewards and success rates. Additionally, DDPG+HER demonstrated increased instability across evaluation episodes with a consistent drop in performance in the later episodes.

E. Observations and discussions

Goal distribution heterogeneity introduced observation skewness across clients, impacting learning and resulting in reduced performance. In the IID federated setting, varying exploration noise across clients in DDPG + HER improved performance by enabling diverse exploration across robots. SAC consistently demonstrated more stable training and better performance than DDPG, indicating its suitability for continuous control tasks.

V. Conclusion

These early observations highlight gaps in training performance, indicating the need for improvements in the training strategies and motivate research towards exploration of methods to improve robot learning under non-IID. Building on this, our future work will be towards extending the present study to more complex manipulation tasks (e.g., grasping), exploring model and data-based approaches to address non-IID in federated robot learning, and analyzing the trade-off between device scalability and model performance in federated settings.

References

[1] O. Zaland, E. Elmroth, and M. Bhuyan, "MTF-Grasp: A Multi-tier Federated Learning Approach for Robotic Grasping," Jul. 16, 2025, *arXiv*: arXiv:2507.10158. doi: 10.48550/arXiv.2507.10158.

[2] B. Sun, X. Song, Y. Tu, and M. Liu, "FedAgent: Federated Learning on Non-IID Data via Reinforcement Learning and Knowledge Distillation," *Expert Syst. Appl.*, vol. 285, p. 127973, Aug. 2025, doi: 10.1016/j.eswa.2025.127973.

[3] G. Liao, M. Jogan, E. Eaton, and D. A. Hashimoto, "FORLA: Federated Object-centric Representation Learning with Slot Attention," Oct. 24, 2025, *arXiv*: arXiv:2506.02964. doi: 10.48550/arXiv.2506.02964.

[4] S. B. Betran, A. Longhini, M. Vasco, Y. Zhang, and D. Kragic, "FLAME: A Federated Learning Benchmark for Robotic Manipulation," Mar. 03, 2025, *arXiv*: arXiv:2503.01729. doi: 10.48550/arXiv.2503.01729.

[5] C. Miao *et al.*, "FedVLA: Federated Vision-Language-Action Learning with Dual Gating Mixture-of-Experts for Robotic Manipulation," Aug. 04, 2025, *arXiv*: arXiv:2508.02190. doi: 10.48550/arXiv.2508.02190.

[6] X. Ma, J. Zhu, Z. Lin, S. Chen, and Y. Qin, "A state-of-the-art survey on solving non-IID data in Federated Learning," *Future Gener. Comput. Syst.*, vol. 135, pp. 244–258, Oct. 2022, doi: 10.1016/j.future.2022.05.003.

[7] O. Zaland, C. Nguyen, F. T. Pokorny, and M. Bhuyan, "Federated Learning for Large-Scale Cloud Robotic Manipulation: Opportunities and Challenges," Jul. 23, 2025, *arXiv*: arXiv:2507.17903. doi: 10.48550/arXiv.2507.17903.

[8] X. Zhu, F. Zhang, and H. Li, "Swarm Deep Reinforcement Learning for Robotic Manipulation," *Procedia Comput. Sci.*, vol. 198, pp. 472–479, 2022, doi: 10.1016/j.procs.2021.12.272.

[9] M. Plappert *et al.*, "Multi-Goal Reinforcement Learning: Challenging Robotics Environments and Request for Research," Mar. 10, 2018, *arXiv*: arXiv:1802.09464. doi: 10.48550/arXiv.1802.09464.

[10] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor," 2018, *arXiv*. doi: 10.48550/ARXIV.1801.01290.

[11] T. P. Lillicrap *et al.*, "CONTINUOUS CONTROL WITH DEEP REINFORCEMENT LEARNING," 2016.

[12] M. Andrychowicz *et al.*, "Hindsight Experience Replay," Feb. 23, 2018, *arXiv*: arXiv:1707.01495. doi: 10.48550/arXiv.1707.01495.

[13] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," 2016, doi: 10.48550/ARXIV.1602.05629.