Poster: FedRos -Federated Reinforcement Learning for Networked Mobile-Robot Collaboration

4 authors:


Wenli Xiao
Carnegie Mellon University


Tingwei Ye
New York University


Bing Luo
Duke Kunshan University


Jianwei Huang
The Chinese University of Hong Kong

Poster: FedRos - Federated Reinforcement Learning for Networked Mobile-Robot Collaboration

Wenli Xiao[*†], Tingwei Ye[‡], Bing Luo[§], Jianwei Huang[*†]
[*]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China
[†]Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China
[‡]Tandon School of Engineering, New York University, New York, USA
[§]Electrical and Computer Engineering, Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, China
Email: wenlixiao@link.cuhk.edu.cn, ty2281@nyu.edu, bing.luo@dukekunshan.edu.cn, jianweihuang@cuhk.edu.cn

***Abstract*—In this paper, we propose FedRos, a Federated Reinforcement Learning based multi-robot system, which enables networked robots collaboratively to train a shared model without sharing their private sensing data. Firstly, we present the FedRos pipeline that embeds the Webots robotics simulator. We then highlight features of FedRos, including its compatibility with the state-of-the-art Federated Learning and Reinforcement Learning algorithms and its sim-to-real viability. Lastly, we present benchmark experiments to show the effectiveness of FedRos.**[1]

***Index Terms*—Federated Learning, Robotics Networks**

I. Introduction

Navigation in unknown environments is one of the most crucial problems in robotics. Recent research in mobile navigation has achieved significant performance improvement by employing Deep Reinforcement Learning (DRL). Compared to single robot deployment, multi-robot DRL frameworks are utilized to improve the exploring efficiency.

However, most existing multi-robot frameworks require sending the raw sensing data back to the central controller for centralized model training, which causes severe communication delay that renders real-time applications impractical. Moreover, transmitting these raw data may also cause privacy leakage if the exploring environment is sensitive.

To tackle these challenges, we propose FedRos, an end-to-end collaborative robotics learning framework that leverages Federated Learning (FL). In our framework, only the locally trained model parameters are iteratively exchanged between robots and the central server, which not only improves communication efficiency but also protects data privacy. Therefore, the total learning efficiency can be significantly improved as robots acquire knowledge from others, even if in environments it has never seen before.

We implement the FedRos system using a simulated robotics platform, where we conduct a benchmark experiment in mobile navigation. The details of the simulation setup and the results of the experiment are discussed in Sec. III. The purpose of the experiment is to showcase the ability of robots to navigate successfully in various environments through federated learning. Our work represents a significant first step in the application of FL in multi-robot collaboration. We plan to extend our research by implementing FedRos in real-world robotics systems in the future.

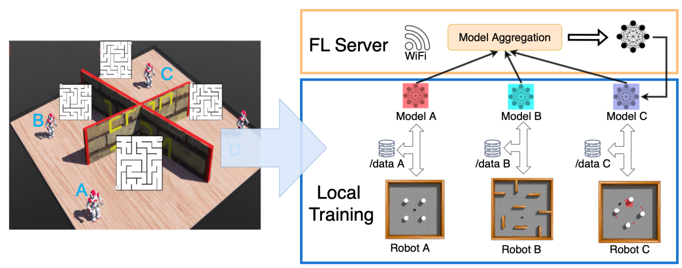

Fig. 1. A scenario of Federated Deep Reinforcement Learning in navigation tasks (left) and FedRos pipeline (right).

II. Design Overview

The left side of Figure 1 illustrates the concept of FedRos enabling a group of robots to explore an unfamiliar area together, with each robot focusing on a specific subarea. In each round, the robots independently apply DRL algorithms to their collected sensor data and upload their trained models to the server. The server then combines these models using weighted aggregation and distributes the global model to the robots for the next training round. After the FL process is completed, the robots are able to achieve a comprehensive model of the entire area.

We construct FedRos based on the Federated Learning pipeline, including 1) a FL server for model aggregation and 2) many mobile robots[2] run DRL algorithms simultaneously. We highlight three aspects in the following sections:

1) In Sec. II-A, we introduce the pipeline of FedRos that enables the full-stack federated learning process simulated on one computation machine to utilize the multi-core processors.

[1]Jianwei Huang and Bing Luo are co-corresponding authors of this paper. This work is supported by the National Natural Science Foundation of China (Project 62271434), Shenzhen Science and Technology Program (Project JCYJ20210324120011032), Guangdong Basic and Applied Basic Research Foundation (Project 2021B1515120008), Shenzhen Key Lab of Crowd Intelligence Empowered Low-Carbon Energy Network (No. ZDSYS20220606100601002), and the Shenzhen Institute of Artificial Intelligence and Robotics for Society. Video of FedRos demo: https://youtu.be/oDVpB6eo6qs

[2]Robots may be heterogeneous, e.g., from different companies.

2) In Sec. II-B, we present the DRL wrap-up of FedRos, which piggybacks on Gym [1] and supports quickly deploying the state-of-the-art DRL algorithms.
3) In Sec. II-C, we show the viability of deploying FedRos to the real-world robotics system in the future.

A. Federated Learning Pipeline

The right side of Figure 1 illustrates the federated learning pipeline of FedRos. FedRos simulates FL via creating two different CPU processes: the FL Server and Local Training. These processes run sequentially until obtaining a considerable global model.

FL Server process can run different model aggregation algorithms, and we apply FedAvg [2] as the default. We implemented model transition by reserving a shared file directory on disk to store robots' local models, where the aggregation algorithm access to fetch the weights of models.

Local Training process creates multiple CPU threads to simulate local DRL training for multiple robots. FedRos creates a private file directory for each thread to store the robot's private perception data. To align with the Federated Learning setting, we enforce that private data is only accessible to its owner.

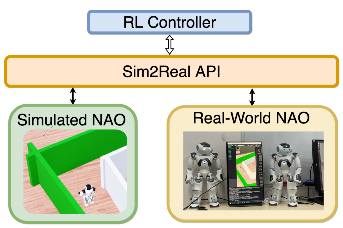

Fig. 2. RL controllers can be agnostic to simulation and real-world environments using Sim2Real API.

B. Reinforcement Learning Wrap-up

We implement the DRL environment on Webots, which is a professional software for simulating mobile robots, and simulate the NAO[3] robot as the DRL agent. For robots' local learning, we consider the standard RL settings:

- *State* S. We choose Sonar and GPS signals as the observation state, which enable robots to have sufficient perception for navigation. In our setup, we did not use the visual and auditory information[4].
- *Actions* A. We utilize Webots protocols to implement the locomotion controller at the bottom level, which enables the DRL agent to command the waypoint.
- *Reward function* R. We derive a reward function for robot navigation tasks. Namely, the reward motivates robots to move to the target and avoid collisions.
- *Robot Reset.* During DRL training, humanoid NAO robots sometimes fall or become stuck, which causes havoc. To mitigate this problem, FedRos monitors robot behaviors in real-time and resets on failure.

[3]NAO is a functional humanoid robot for perception and locomotion tasks.
[4]We keep the API for fetching visual and audio data, which could be used for multi-modal Reinforcement Learning.

On top of that, we adopt Deep Deterministic Policy Gradient (DDPG) to implement the DRL for robots' local training. Since we implement the RL environment based on Gym API, it is effortless to deploy other state-of-the-art RL algorithms.

C. Sim-To-Real Viability

We consider conducting real-world experiments in the future. To seamlessly transfer FedRos to real-world robotics systems, we develop the Sim2Real API (Fig. 2) to align the controller of real-world NAO with the simulation. Thus, we only need minor modifications to integrate real-world NAO robots into the FedRos training loop.

III. Experiment

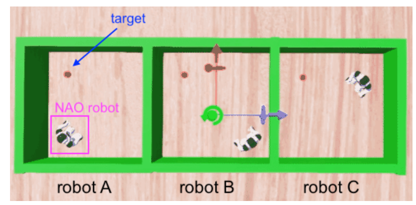

Fig. 3. The experiment is simulated in Webots [3]. In each room, the green square denotes the obstacle, and the brown dot point is the target.

A. Experiment setup

We demonstrate the FedRos prototype on a simulated benchmark task. Figure 3 shows the experiment where three NAO robots cooperatively learn to navigate from scratch in different rooms.

B. Networked Mobile Navigation

The demonstration will feature a collaborative learning experience using three NAO robots. The robots will work together to learn how to navigate from scratch in three rooms, as shown in Figure 3. The target location is fixed, and the robots are placed in different positions within each room. The goal is to demonstrate how FedRos can create a global model through federated learning, enabling any robot to successfully navigate to the target from any starting position. Figure 4 shows the federated learning policy eventually converges.

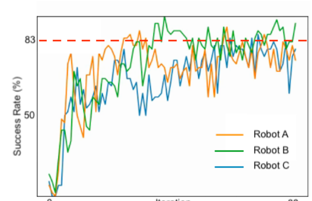

Fig. 4. Goal reaching success rate of three robots versus the training iterations.

References

[1] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," *arXiv preprint arXiv:1606.01540*, 2016.
[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.
[3] O. Michel, "Cyberbotics ltd. webots™: professional mobile robot simulation," *International Journal of Advanced Robotic Systems*, vol. 1, no. 1, p. 5, 2004.