Loss-Dynamics-Aware Multi-Agent Cooperative Active Sampling System for Efficient Robotic Online Learning

Anonymous Author(s)
Submission Id: 320*

Abstract

Online training tasks rely on training input consecutively sampled from the real world to refine their training model, and automated mobile devices (robots) shouldering the responsibility of sampling has the potential to further enhance their training performance by actively sampling in reaction to the real-time training quality, besides saving human labor. However, to achieve such active sampling faces two major gaps. First, as the training quality is typically computation-intensive to get for a robot (e.g., validating), where to move for the robot to boost training performance? Second, since the consecutive samples from nearby states (e.g., position, orientation) of a robot are too similar and would limit the training information gain, how to move for better efficiency?

We observe that real-time training loss is spatially and temporally related to locations in the environment and implies the level of information gain on further sampling of these locations. The second gap can be overcome by cooperative multi-view sampling from multiple robots, since they are naturally distant in state space. Based on these observation, we choose to build an online training application named MIAS (Multi-robot Implicit Active SLAM) that drives multiple robots to actively and cooperatively sample the environment in real-time reaction to the quality of the training implicit SLAM model. Evaluation shows that MIAS with three robots speeds up the implicit SLAM tasks not only by up to xxX compared to the baselines with three robots, but also by xxX compared to MIAS with single robot.

CCS Concepts: • **Computer systems organization → Robotic control**.

Keywords: Online Training, Active Learning, Active Sampling, Multi-Agent System, 3D Reconstruction

1 Introduction

The recent flourishing of machine learning is emphasizing the importance of both quantity and quality (or information gain) of training data for high performance (e.g., accuracy) of an Artificial Intelligence (AI) model. For online training tasks typically (e.g., domain adaptation, implicit rendering, long term learning) deployed over mobile robots which consecutively take unlabelled sample (e.g., images, lidar) from the environment as training input for an AI model, enabling automatic acquisition of high quality training data on the robots will boost the performance of their training AI model.

Unfortunately, although the existing related domain, active learning, aims to find possible training data of highest quality from the environment for the training AI model, the found training data as the acquisition target fails to form a navigation path for the robot to collect high quality training data. Specifically, the robot is moving and sampling in a continuous world and along the path the robot tends to collect similar samples as the acquisition target, which in turn causes more low quality training data being input. In our evaluation, such methods cause the similarity between consecutively collected training data to be over xx% (measured in SSIM) and thus low information gain. Another problem of such methods is that such methods cannot be scaled to multiple robots, since TODO.

The key reason of the above problem is that when searching for navigation path (the acquisition destination), they are viewing the training AI model in a static perspective. As online training of the AI model proceeds, the parameters of the AI model keeps evolving and thus, the possible training data of highest quality is also changed (i.e., AI model state evolves), especially when the robot starts navigating and new training data is collected. Failure to model such AI model state evolution in path searching leads to low quality of training data collected along the navigation path.

To tackle the above problem, a dynamic perspective of the training AI model is necessary while searching for a path, which is difficult to predict since the training data sampled in the future are not available. Instead, we observe that if we relate the training loss with the zones of the environment of their corresponding training input, its dynamics have the potential to predict future training loss if the same zone is sampled in the future, whose diff reflects possible information gain and AI model state evolution. The predicted future training loss can recursively serve as the basis of the prediction of the training loss when a further step is taken and finally forms the prediction of the accumulated information gain along the whole path, with which we can search for an optimal path with highest accumulated information gain.

Notably, such method can easily be extended to multi-agent cooperation: an agent can directly plan its own path based on the predicted training loss from the planned path shared by other agents, which enables co-operative multi-view sampling across different agents, so as to alleviate the limitation of the continuous action and sampling space of a single robot.

Based on these ideas, we propose LOss-Dynamics-Aware multi-agent cooperative active sampling system for robotic online training (LODA). In LODA, we distribute the online training workload across all agents involved and renders the training loss to a sparse 3D grid on each agent which acts as the representation of the training loss dynamics. With this representation, we propose LOss-Dynamics-Aware-Random Rapid Tree star (LODA-RRT*) algorithm: we search for candidate paths based on RRT* algorithm and use a small multi-layer perception model (MLP) to online learn and recursively predict the future training loss after taking each step of a candidate path, which approximates the accumulated information gain along the path. If planned paths from other robots are received, their predicted future training loss will be temporarily merged into the grid representation as the basis of path planning. In this way, each agent finds a cooperative multi-view sampling path that optimizes accumulated information gain by considering the AI model state evolution modelled by loss dynamics, leading to higher quality of collected training data and higher performance of the training AI model.

We implemented LODA based on ROS2 Galactic and Python3.8 on Ubuntu20.04 and evaluated its performance across one to three robots involved. We choose a state-of-the-art (SOTA) online training task BNV-Fusion in the domain of implicit rendering as the major workload and use Replica, a standard computer-vision dataset, together with Habitat simulator as the dataset and evaluated across different level of sizes of the environments. We select two SOTA active learning baselines, namely Badge and SDF as the major baselines. Evaluation shows that:

- LODA is efficient. Under the same number of robots, same test scene and same sampling and training duration, LODA improved the accuracy of the training model by xx% ~ xx% compared with the baselines due to the improved quality of sampled training input from the environment.
- LODA is scalable. When the number of robots increased from one to three, the advantage in accuracy of LODA increased from xx% ~ xx% to xx% ~ xx% compared with the baselines. When the scale of the environment increased, the advantage in accuracy of LODA also increased from xx% ~ xx% to xx% ~ xx% compared with the baselines.
- LODA is cooperative. Ablation study shows that the cooperation by sharing the planned path among all robots contributed to xx% ~ xx% of the accuracy improvement compared with the cases where the cooperation is disabled.

The major contributions of this paper is that we propose a novel dynamic perspective of training AI model modelled with the loss dynamics of the environment for automatic high quality training data acquisition in robotic online training tasks to improve the performance of training AI model. Under this idea, our proposed system LODA and the LODA-RRT* algorithm together model the accumulated training information gain of candidate paths and find an optimal path for robot sampling that optimizes the overall quality of collected training data. It also enables efficient multi-robot multi-view cooperation that further increases the quality of training data by simply sharing the planned path among the robots. LODA and LODA-RRT* will nurture the development of more real-world robotic online training applications and improving their performance by enabling efficient and high performance automatic high quality training data acquisition.

2 Background & RELATED WORK

2.1 Online Training

2.2 Active Learning

2.3 Active Sampling in Traditional Robotics

3 Overview

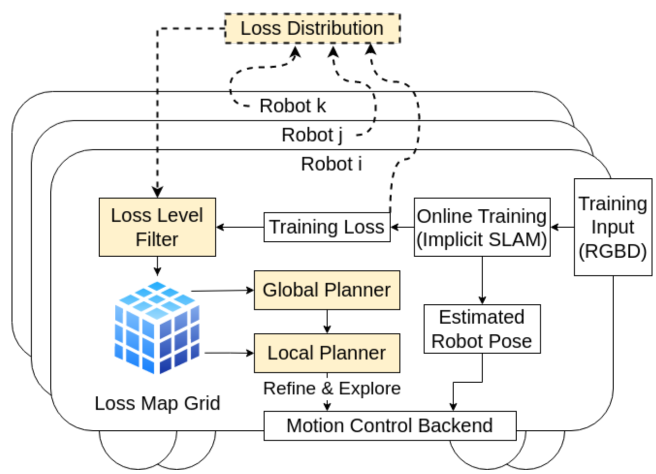

Figure 1. Overview of MIAS

This chapter presents the architecture of LODA and gives an overview of how LODA models the loss dynamics of the environment and how it is used to search for path with optimal accumulated information gain with a single robot or the cooperation of multiple agents in the LODA-RRT* algorithm. Assume that each robot is a four-wheel robot with sensors (e.g., cameras, lidar) attached to it and it is periodically taking samples from the sensors as the training input for an underlying online training task (e.g., 3D reconstruction).

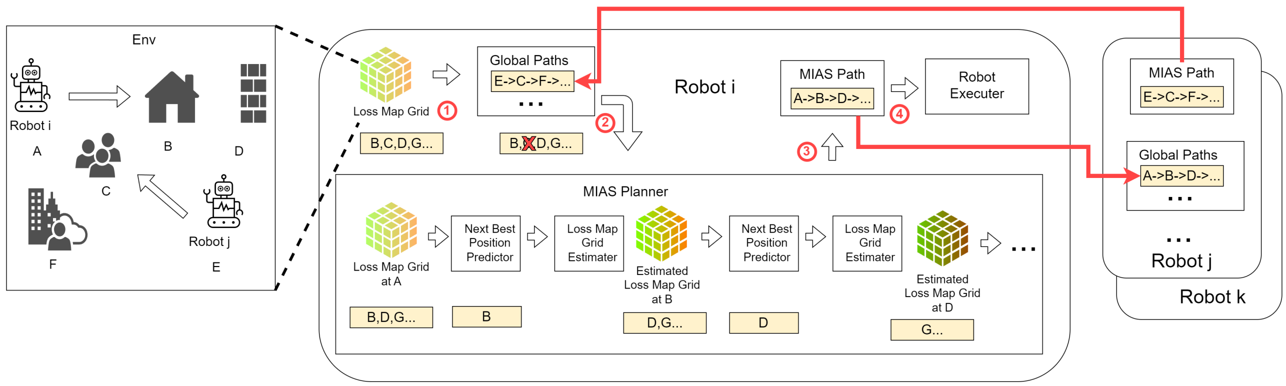

Figure 2. New Overview of MIAS

3.1 Loss Rendering

3.2 Loss Dynamics Modelling

3.3 Future Loss Prediction

3.4 Recursive Path Planning

4 Design

4.1 Loss Dynamics Modelling

4.2 LODA-RRT* Algorithm

5 Evaluation

baselines: badge, uncertainty, random

5.1 End-to-End Performance

5.1.1 Different Number of Robots. figures: y accuracy; x time; 1 robot; 2 robots; 3 robots

fact

ours all better than the baselines; more robot, diff larger badge, uncertainty > random, but due to ... with our system, ...

5.1.2 Different Size of Scene. figure: y accuracy; x time; small scene; medium scene; large scene

facts

ours advantages remain, larger scene, better badge, uncertainty > random, but due to ...

our can be better because...

5.2 Ablation Study

stop multi robot cooperation in our system, get worse performance

References

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009