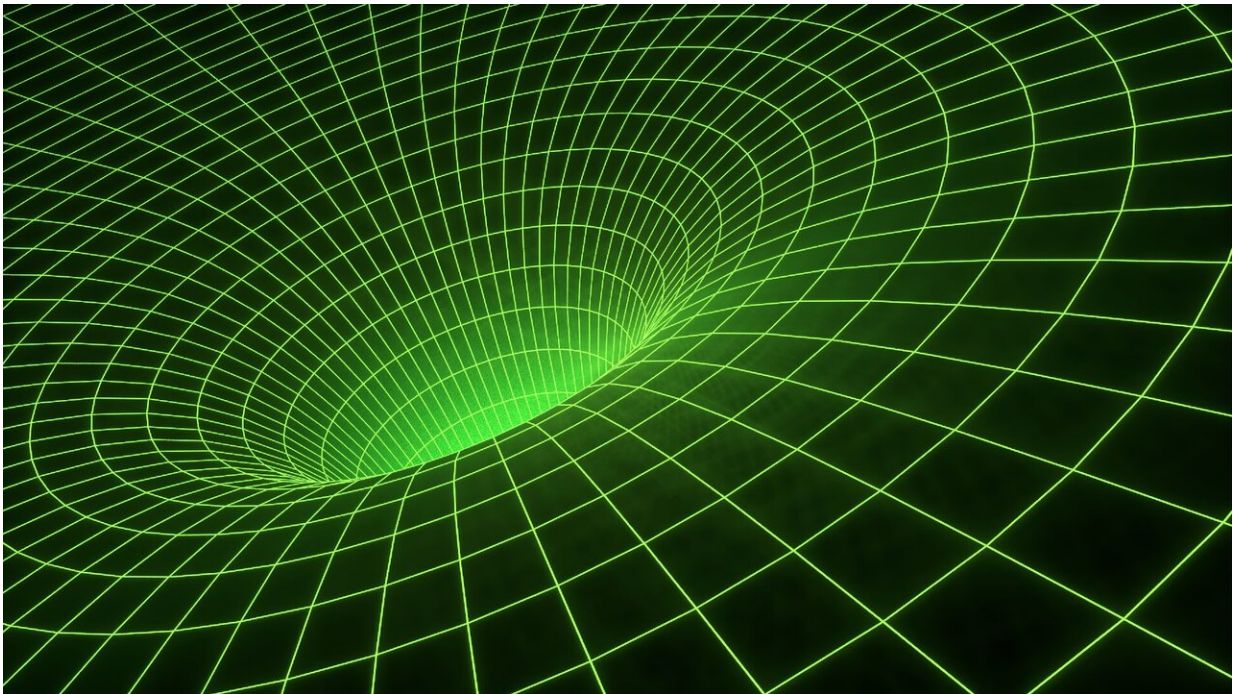


Researchers introduce new algorithm to reduce machine learning time

March 9 2020, by Li Yuan



Credit: CC0 Public Domain

A research team led by Prof. LI Huiyun from the Shenzhen Institutes of Advanced Technology (SIAT) of the Chinese Academy of Sciences introduced a simple deep reinforcement learning (DRL) algorithm with m-out-of-n bootstrap technique and aggregated multiple deep deterministic policy gradient (DDPG) algorithm structures.

Named "bootstrapped aggregated multi-DDPG" (BAMDDPG), the new [algorithm](#) accelerated the training process and increased the performance in the area of intelligent artificial research.

The researchers tested their algorithm on 2-D robot and open racing car simulator (TORCS). The experiment results on the 2-D robot arm game showed that the reward gained by the aggregated policy was 10%-50% better than those gained by subpolicies, and experiment results on the TORCS demonstrated that the [new algorithm](#) could learn successful control policies with less training time by 56.7%.

DDPG algorithm operating over continuous space of actions has attracted great attention for [reinforcement learning](#). However, the exploration strategy through dynamic programming within the Bayesian belief state space is rather inefficient even for simple systems. This usually results in failure of the standard bootstrap when learning an optimal [policy](#).

The proposed algorithm uses the centralized experience replay buffer to improve the exploration efficiency. M-out-of-n bootstrap with random initialization produces reasonable uncertainty estimates at low computational cost, helping in the convergence of the training. The proposed bootstrapped and aggregated DDPG can reduce the learning time.

BAMDDPG enables each agent to use experiences encountered by other agents. This makes the training of subpolicies of BAMDDPG more efficient since each agent owns a wider vision and more environment information.

This method is effective to the sequential and iterative training data, where the data exhibit long-tailed distribution, rather than the norm distribution implicated by the independent identically distributed data

assumption. It can learn the optimal policies with much less [training](#) time for tasks with continuous space of actions and states.

The study, titled "Deep Ensemble Reinforcement Learning with Multiple Deep Deterministic Policy Gradient Algorithm," was published in *Hindawi*.

More information: Junta Wu et al. Deep Ensemble Reinforcement Learning with Multiple Deep Deterministic Policy Gradient Algorithm, *Mathematical Problems in Engineering* (2020). [DOI: 10.1155/2020/4275623](#)

Provided by Chinese Academy of Sciences

Citation: Researchers introduce new algorithm to reduce machine learning time (2020, March 9) retrieved 26 April 2024 from <https://techxplore.com/news/2020-03-algorithm-machine.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.