2022 Sixth IEEE International Conference on Robotic Computing (IRC) | 978-1-6654-7260-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/IRC55401.2022.00022

GA+DDPG+HER: Genetic Algorithm-Based Function Optimizer in Deep Reinforcement Learning for Robotic Manipulation Tasks

Adarsh Sehgal*†, Nicholas Ward*†, Hung Manh La*†, Christos Papachristos†, and Sushil Louis†
* Advanced Robotics and Automation (ARA) Laboratory
† Department of Computer Science and Engineering, University of Nevada, Reno, 89557, NV, USA

***Abstract*—Agents can base decisions made using reinforcement learning (RL) on a reward function. The selection of values for the learning algorithm parameters can, nevertheless, have a substantial impact on the overall learning process. In order to discover values for the learning parameters that are close to optimal, we extended our previously proposed genetic algorithm-based Deep Deterministic Policy Gradient and Hindsight Experience Replay approach (referred to as GA+DDPG+HER) in this study. On the robotic manipulation tasks of FetchReach, FetchSlide, FetchPush, FetchPick&Place, and DoorOpening, we applied the GA+DDPG+HER methodology. Our technique GA+DDPG+HER was also used in the AuboReach environment with a few adjustments. Our experimental analysis demonstrates that our method produces performance that is noticeably better and occurs faster than the original algorithm. We also offer proof that GA+DDPG+HER beat the current approaches. The final results support our assertion and offer sufficient proof that automating the parameter tuning procedure is crucial and does cut down learning time by as much as 57%.**

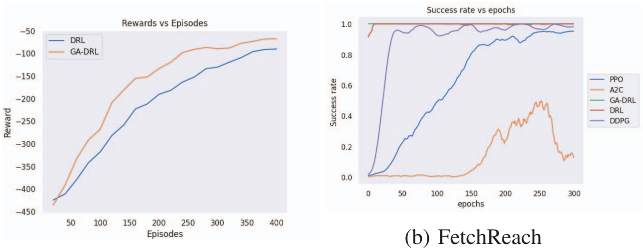
***Index Terms*—DRL, DDPG+HER, Reinforcement Learning, Genetic Algorithm, GA+DDPG+HER, DDPG, HER**

I. Introduction

Recently, Reinforcement Learning (RL) has been put to significant uses such as standard robotic manipulation [1]. Each of these applications make use of RL as an encouraging substitute to automating manual effort.

In this paper, we are specifically using DDPG [2] combined with HER [3] to train Deep Reinforcement Learning (DRL) policies. Our primary contributions are as follows: 1) For algorithm analysis, we deployed our own previously developed GA+DDPG+HER algorithm [4] on Aubo-i5 simulated and real environments. 2) As the GA develops, the training process is examined utilizing a variety of criteria. 3) GA+DDPG+HER was compared to existing techniques. Open source code is available at https://github.com/aralab-unr/ga-drl-aubo-ara-lab.

The algorithm GA+DDPG+HER was discovered in our earlier work [4] and [5]. [6] provides a thorough background, experimental conditions, implementation details, training evaluation, and analysis of this research. Some of the closely related work includes [7]. The results from these papers provide more evidence that when a GA is used to automatically tune the hyper-parameters for DDPG+HER, efficiency can be largely improved. The difference can greatly impact the time it takes for a learning agent to learn.



(a) AuboReach

(b) FetchReach

Fig. 1: GA+DDPG+HER vs.. existing algorithm's efficiency evaluation plots (total reward vs.. episodes) when all the six parameters are found by GA. FetchReach is averaged over two runs, and AuboReach is averaged over ten runs.

II. Experimental Results

A. Training evaluation

Figure 1 contrasts DDPG+HER with GA+DDPG+HER, whereas Figure 2 shows how the system performs better as the GA evolves.

B. Efficiency evaluation

This section compares DDPG+HER to GA+DDPG+HER in a variety of environmental settings. The most effective algorithm is indicated in each of the tables I-IV by a number in bold.

Method	FetchPick&Place	FetchPush	FetchReach	FetchSlide	DoorOpening	AuboReach
DDPG+HER	6,000	2,760	100	**4,380**	960	320
GA+DDPG+HER	**2,270**	**1,260**	**60**	6,000	**180**	**228**

TABLE I: Efficiency evaluation: Average (over ten runs) episodes comparison to reach the goal, for all the tasks.

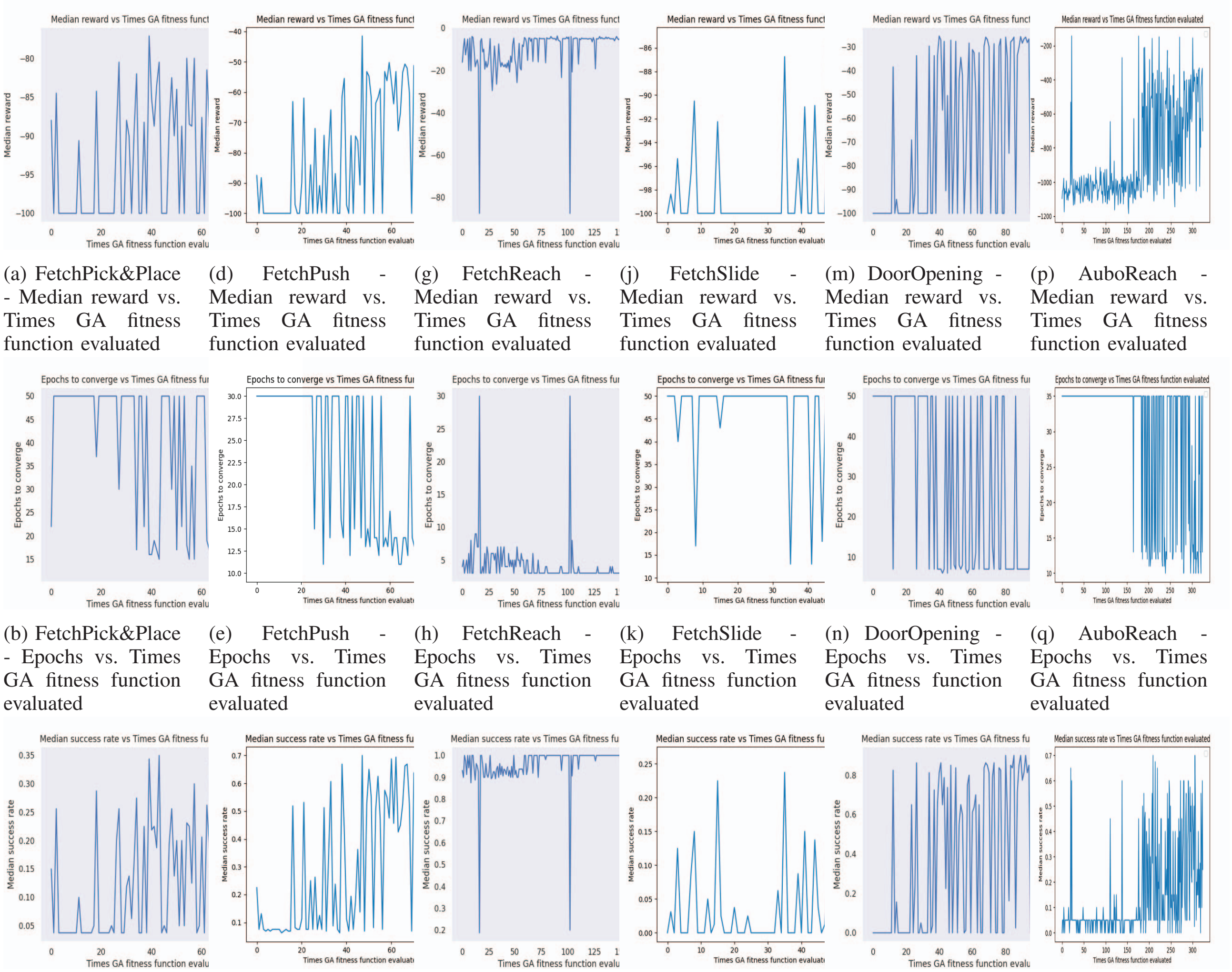
Method	FetchPick&Place	FetchPush	FetchReach	FetchSlide	DoorOpening	AuboReach
DDPG+HER	3069.981	1314.477	47.223	**2012.645**	897.816	93.258
GA+DDPG+HER	**1224.697**	**565.178**	**28.028**	3063.599	**167.883**	**66.818**

TABLE II: Efficiency evaluation: Average (over 10 runs) running time (s) comparison to reach the goal, for all the tasks.

Method	FetchPick&Place	FetchPush	FetchReach	FetchSlide	DoorOpening	AuboReach
DDPG+HER	300,000	138,000	5000	**219,000**	48000	65,600
GA+DDPG+HER	**113,000**	**63,000**	**3000**	300,000	**9000**	**46,000**

TABLE III: Efficiency evaluation: Average (over ten runs) steps comparison to reach the goal, for all the tasks.

978-1-6654-7260-9/22/$31.00 ©2022 IEEE
DOI 10.1109/IRC55401.2022.00022



(a) FetchPick&Place - Median reward vs. Times GA fitness function evaluated

(d) FetchPush - Median reward vs. Times GA fitness function evaluated

(g) FetchReach - Median reward vs. Times GA fitness function evaluated

(j) FetchSlide - Median reward vs. Times GA fitness function evaluated

(m) DoorOpening - Median reward vs. Times GA fitness function evaluated

(p) AuboReach - Median reward vs. Times GA fitness function evaluated

(b) FetchPick&Place - Epochs vs. Times GA fitness function evaluated

(e) FetchPush - Epochs vs. Times GA fitness function evaluated

(h) FetchReach - Epochs vs. Times GA fitness function evaluated

(k) FetchSlide - Epochs vs. Times GA fitness function evaluated

(n) DoorOpening - Epochs vs. Times GA fitness function evaluated

(q) AuboReach - Epochs vs. Times GA fitness function evaluated

(c) FetchPick&Place - Median success rate vs. Times GA fitness function evaluated

(f) FetchPush - Median success rate vs. Times GA fitness function evaluated

(i) FetchReach - Median success rate vs. Times GA fitness function evaluated

(l) FetchSlide - Median success rate vs. Times GA fitness function evaluated

(o) DoorOpening - Median success rate vs. Times GA fitness function evaluated

(r) AuboReach - Median success rate vs. Times GA fitness function evaluated

Fig. 2: GA+DDPG+HER training evaluation plots, when all the 6 parameters are found by GA (this is for one GA run.)

Method	FetchPick&Place	FetchPush	FetchReach	FetchSlide	DoorOpening	AuboReach
DDPG+HER	60	27.6	5	**43.8**	47	16
GA+DDPG+HER	**22.6**	**12.6**	**3**	60	**8**	**11.4**

TABLE IV: Efficiency evaluation: Average (over 10 runs) epochs comparison to reach the goal, for all the tasks.

References

[1] A. Sehgal, M. Sehgal, and H. M. La, "Aacher: Assorted actor-critic deep reinforcement learning with hindsight experience replay," *arXiv preprint arXiv:2210.12892*, 2022.

[2] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[3] M. Andrychowicz, F. Wolski, A. Ray, J. Schneider, R. Fong, P. Welinder, B. McGrew, J. Tobin, O. P. Abbeel, and W. Zaremba, "Hindsight experience replay," in *Advances in Neural Information Processing Systems*, 2017, pp. 5048–5058.

[4] A. Sehgal, H. La, S. Louis, and H. Nguyen, "Deep reinforcement learning using genetic algorithm for parameter optimization," in *2019 Third IEEE Int. Conf. on Robotic Computing (IRC)*. IEEE, 2019, pp. 596–601.

[5] A. Sehgal, N. Ward, H. La, and S. Louis, "Automatic parameter optimization using genetic algorithm in deep reinforcement learning for robotic manipulation tasks," *arXiv preprint arXiv:2204.03656*, 2022.

[6] A. Sehgal, N. Ward, H. La, C. Papachristos, and S. Louis, "Deep reinforcement learning for robotic manipulation tasks using a genetic algorithm-based function optimizer," *Encyclopedia with Semantic Computing and Robotic Intelligence*, 2022.

[7] A. Sehgal, A. Singandhupe, H. M. La, A. Tavakkoli, and S. J. Louis, "Lidar-monocular visual odometry with genetic algorithm for parameter optimization," in *Advances in Visual Computing*, G. Bebis *et al.*, Eds. Cham: Springer International Publishing, 2019, pp. 358–370.