Supplementary material for "MCLER: Multi-critic Continual Learning with Experience Replay for Quadruped Gait Generation"

Maoqi Liu*, Yanyun Chen*, Ran Song, Xing Fang, Wenhao Tan, Yibin Li and Wei Zhang

This supplementary material provides extra details for the implementation of MCLER and additional experimental results. Supplementary video demonstrations can be accessed via the following link: https://vsislab.github.io/mcler/.

A. Extra Details of MCLER

Table S1 summarizes the allocation of the sampled and replayed data for ER during the incremental training. As shown in Table S1, the buffer size used to replay for each previously learned gait is $1024*2$, indicating that there is totally $1024*2*(n-1)$ samples of data are are available for replay while training the n_{th} gait, where $n \leqslant 6$.

Table S2 lists the ranges of commands settled during the training procedure for different gait.

TABLE S1

DATA ALLOCATION FOR NORMAL SAMPLE AND EXPERIENCE REPLY

Gait number	Normal	Replay
1	$1024*64$	0
2	$1024*(64-2)$	$1024*2$
...	...	..
N	$1024*(64-2*(n-1))$	$1024*2*(n-1)$

TABLE S2

RANGES OF COMMANDS FOR DIFFERENT GAIT

Gaits	Ranges
Bipedal stand	$\hat{v}_x = 0.0$ m/s, $\hat{v}_y = 0.0$ m/s
Gallop	$\hat{v}_x \in [0.5, 5.0]$ m/s, $\hat{v}_y = 0.0$ m/s
Pace	$\hat{v}_x \in [0.3, 2.0]$ m/s, $\hat{v}_y = 0.0$ m/s
Lateral walk	$\hat{v}_x = 0$ m/s, $\hat{v}_y \in [-1.0, 1.0]$ m/s
Crawl	$\hat{v}_x \in [0.3, 1.0]$ m/s, $\hat{v}_y = 0.0$ m/s
Catwalk	$\hat{v}_x \in [0.5, 1.5]$ m/s, $\hat{v}_y = 0.0$ m/s

B. Additional Experiments

Fig. S1 provides the success rates across the 6 gaits for 10 times with the final policies trained with different methods. We found that the Baseline method maintained a success rate of 0 for the previously learned 5 gaits, indicating that it suffered greatly from catastrophic forgetting during the continual learning. Except for the Baseline method, other methods combined with EWC or ER reached a success rate of 1.0 on the first learned gait (bipedal stand), indicating that EWC and ER could help the robot mitigate catastrophic forgetting on early learned gaits. The methods solely employed with EWC or ER achieved the lower values of success rate for the following learned gaits, reduced from 1.0 to 0 eventually. In contrast, the methods integrated with MC received remarkable success rate across all gaits, in which MCLER obtained the highest success rate of 1.0 over all gaits, demonstrating the superiority of our approach.

Fig. S2 shows the deployment results of the policies trained by different methods. The robot based on the Baseline method performed the catwalk gait, which is the last one it learned, but failed to operate the rest gaits. The robot equipped by MC+ER and MC+EWC were able to behave all of the gaits for a short time yet ultimately failed due to inability to adapting the environments. The robot performed by MCLER successfully produced the expected gaits, including pace, lateral walk, catwalk, and gallop, which demonstrates the robustness and effectiveness of our approach.

Fig. S3 offers the snapshots of the learned bipedal stand gait performed by different methods in simulation. The results shows that the the robot successfully generated the bipedal stand gait except for the one based on the Baseline method, which is consistent with the result shown in Fig. S1.

Maoqi Liu*, Yanyun Chen*, Ran Song, Xing Fang, Wenhao Tan, Yibin Li and Wei Zhang are with the School of Control Science and Engineering, Shandong University, Jinan, China.

Maoqi Liu and Yanyun Chen contributed equally to this work. Corresponding author: Wei Zhang (Email: davidzhang@sdu.edu.cn)

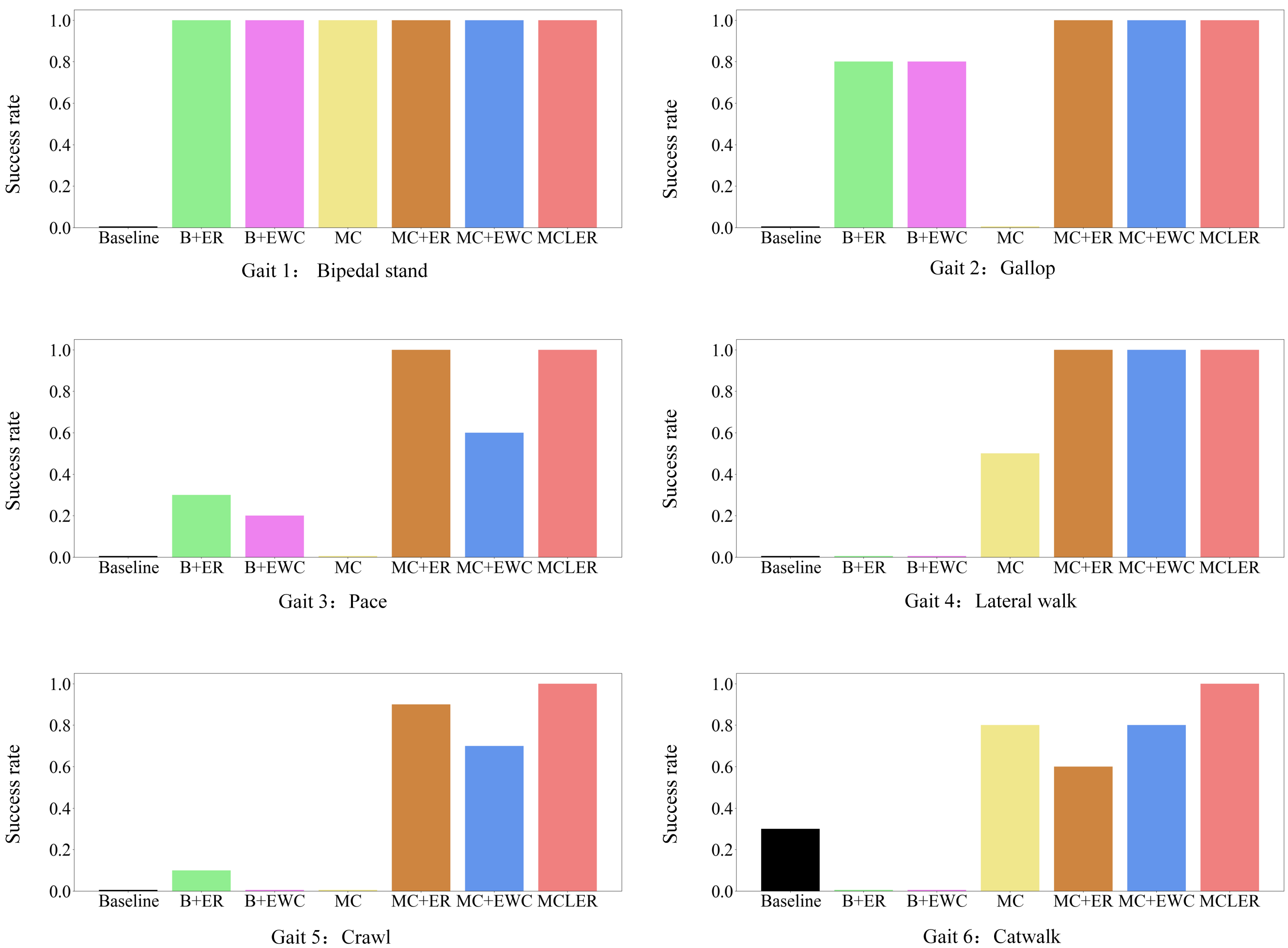

Fig. S1. The evaluated success rate of all gaits learned by different methods at the end of the whole training process.

Fig. S2. The comparison of gaits learned by different methods deployed in the real-world.

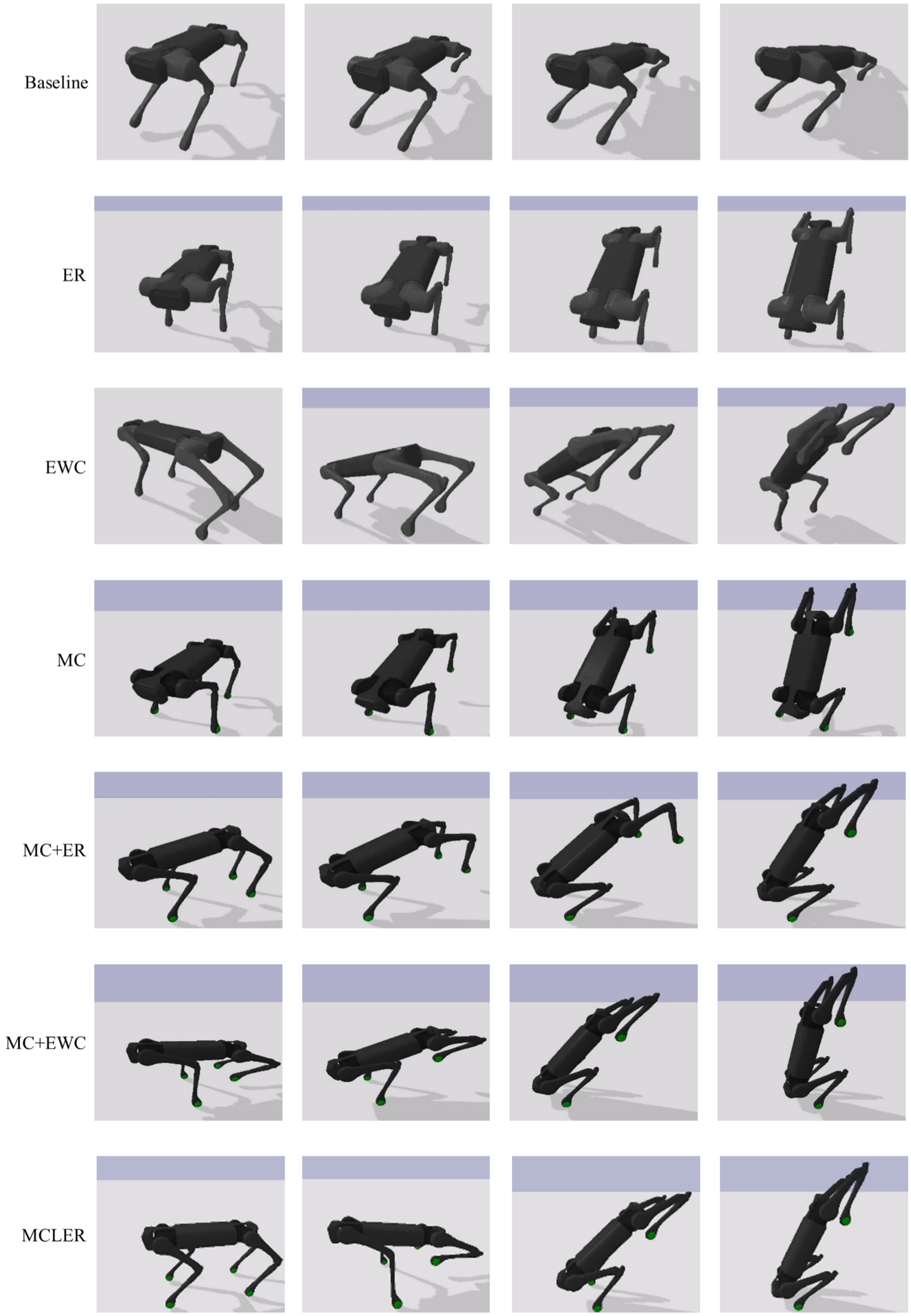

Fig. S3. The comparison of bipedal stand gait learned by different methods shown in the simulation.