

Biological Neurons vs. Deep Reinforcement Learning: Sample efficiency in a simulated game-world



Forough Habibollahi^{1,3*}, Moein Khajehnejad^{2,3*}, Amitesh Gaurav³, Brett J. Kagan³
¹ Dept of Biomedical Engineering, University of Melbourne, ² Dept of Data Science and AI, Monash University,
³ Cortical Labs, Melbourne, Australia
 * Equal Contribution

Abstract

How do synthetic biological systems and artificial neural networks compete in their performance in a game environment?

- We compare the inherent intelligence of *in vitro* biological neuronal networks from human (HCC) and mice (MCC) cells to state-of-the-art deep RL algorithms: DQN, A2C, and PPO in the arcade game 'Pong'.
- We employ DishBrain, a system that embodies *in vitro* neural networks with *in silico* computation using a high-density multielectrode array.
- To account for potential disadvantages occurring as a result of increased input dimensionality to the RL algorithms, we examine 1) IMAGE INPUT as well as two alternative designs of the input structure: 2) PADDLE & BALL POSITION INPUT and 3) BALL POSITION INPUT.
- We find that these very simple biological cultures typically outperform deep RL systems in terms of game performance characteristics implying a higher sample efficiency.

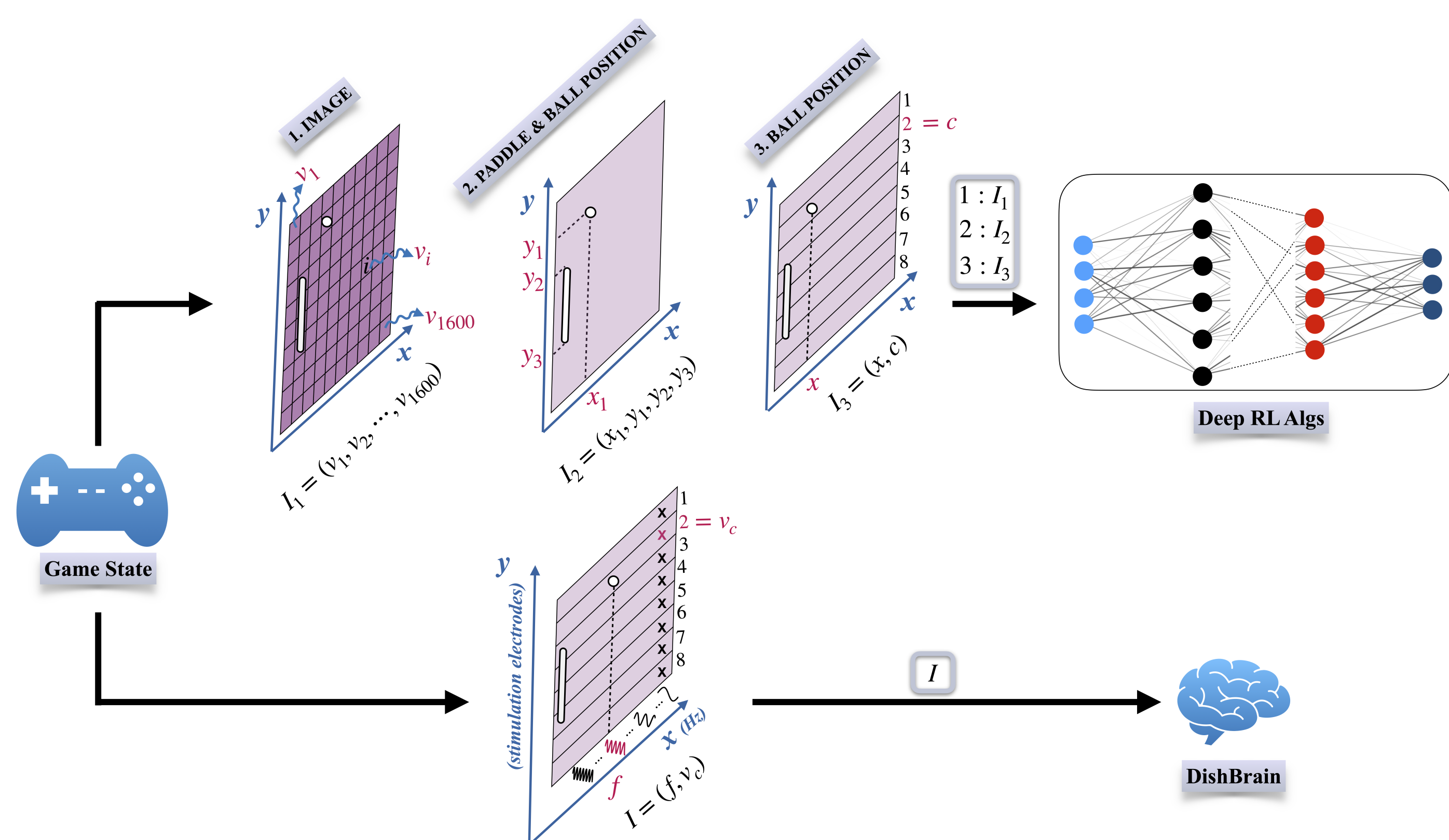


Fig. 1: Information, I , fed in the DishBrain system (bottom) and the three implementations of the deep RL algorithms (top) with inputs of different dimensions.

DishBrain Platform

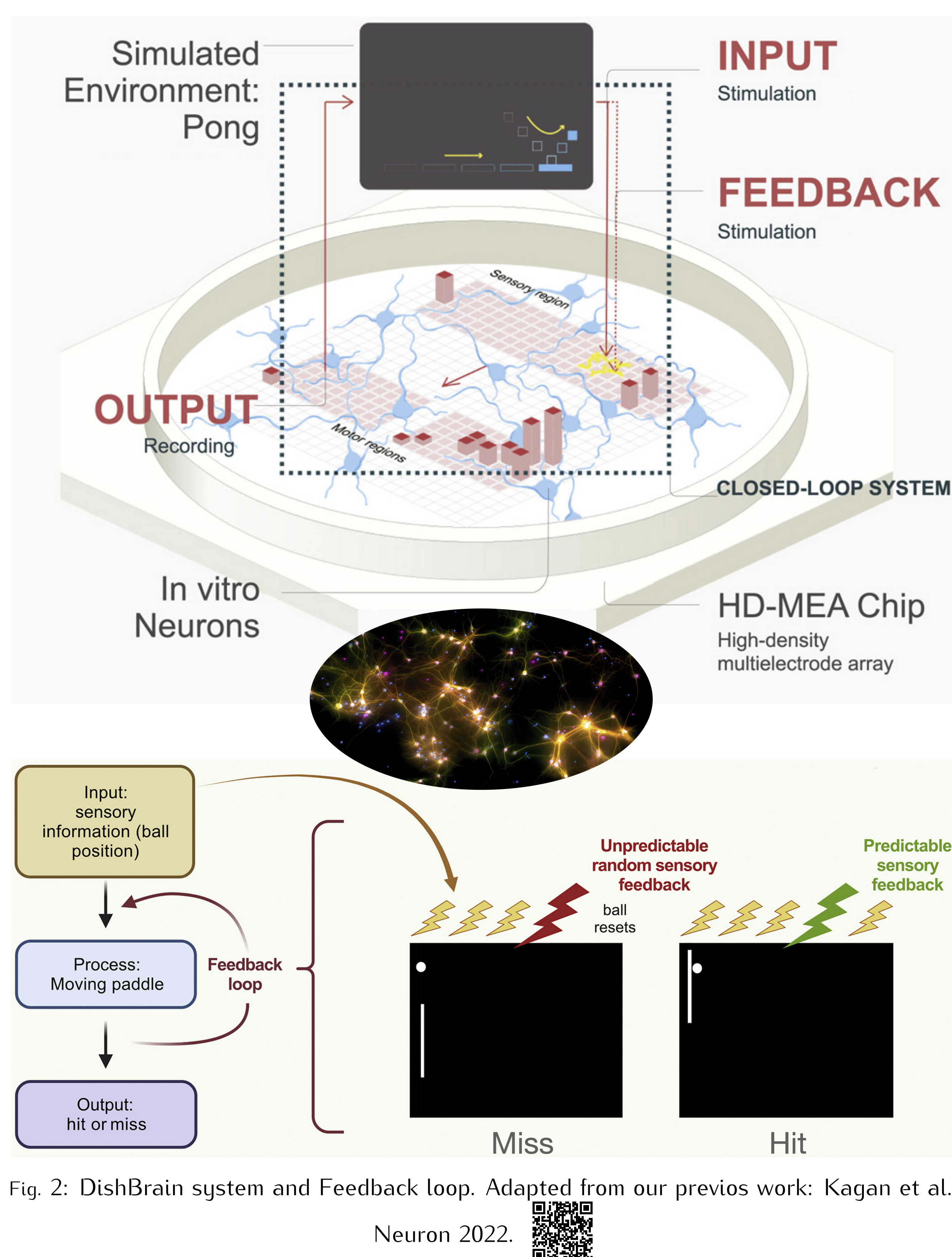


Fig. 2: DishBrain system and Feedback loop. Adapted from our previous work: Kagan et al., Neuron 2022.

Results

- Every 70-episode run of each RL algorithm were first mapped to a real-time equivalent of 20 minutes (length of the biological recordings).
- Key gameplay characteristics including **average rally length**, **% of aces**, and **% of long rallies** (> 3 consecutive hits) were examined for all groups.

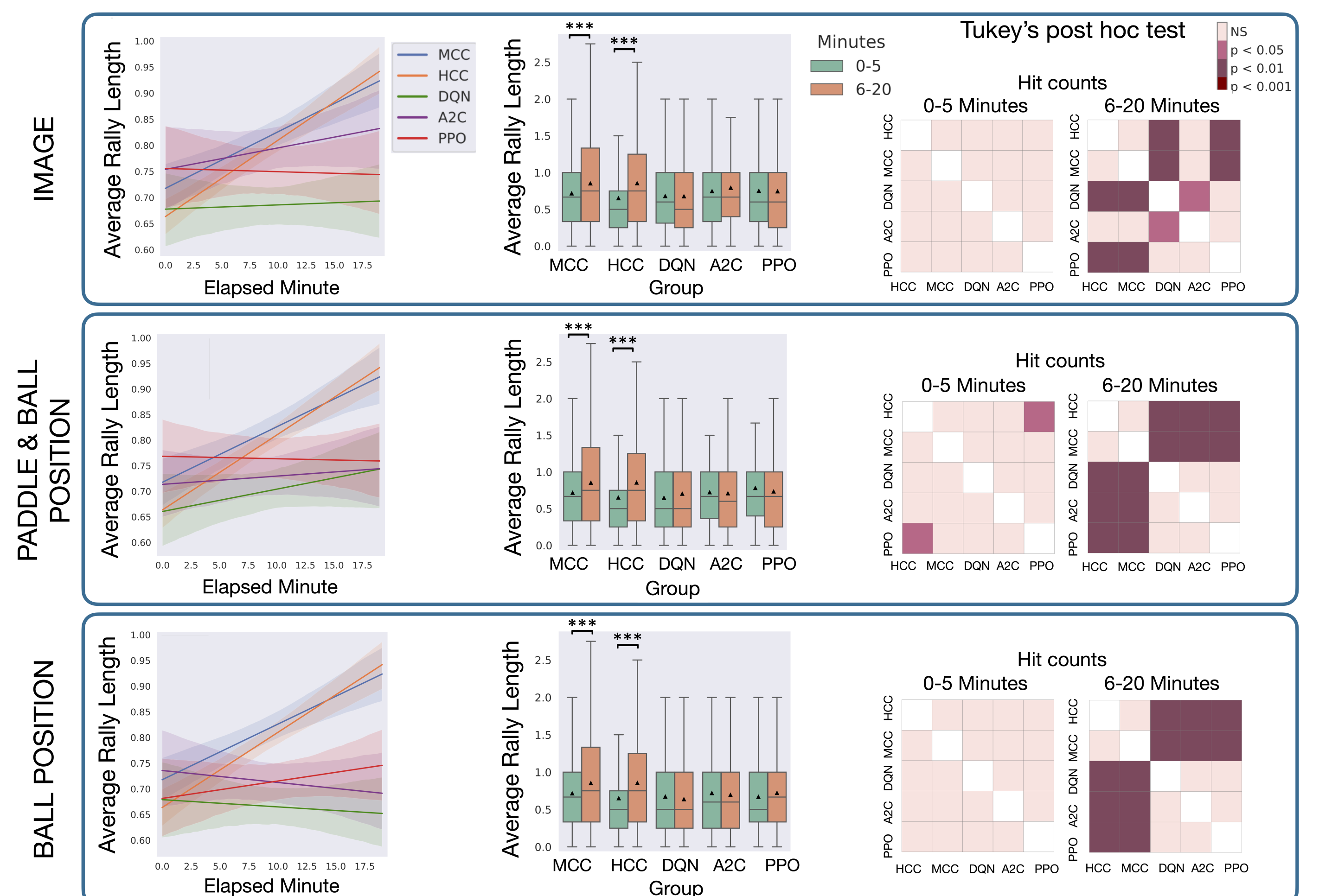


Fig. 3: Comparing game performance in terms of average rally length achieved in different groups. Rows correspond to different designs of the input structure in the RL algorithms.



Fig. 4: LEFT) The average paddle movement in pixels in all the difference groups. RIGHT) Relative improvement (%) in the average hit counts between the first 5 minutes and the last 15 minutes of all sessions.

Discussion

- DQN algorithm was outperformed by all the other groups in all the tests.
- The biological cultures outperformed all RL baselines in terms of the ultimate average rally length, ultimate % of aces (i.e. games lost in a single shot), and ultimate % of long rallies achieved.
- The increase in the average rally length, the decrease in the number of aces, and the increase in the number of long rallies were only significant within the MCC and HCC biological cultures.
- The HCC group had the highest relative improvement in the average number of hits between the first 5 minutes and last 15 minutes of the game.

- ✓ This is the first comparison between a synthetic biological intelligence system and state-of-the-art RL algorithms.
- ✓ This work establishes that even the most rudimentary synthetic biological intelligence systems with limited informational input are a viable learning system that can compete and even outperform the RL algorithms.
- ✓ Coupled with the promise of significant gains in power efficiencies and flexibility of tasks, these synthetic biological intelligence system present a compelling pathway for realizing real-time learning.