frontiers

Supplementary Material

1 Supplementary Figure

Supplementary Figure 1 displays an itemized summary of the reviewed neural decoders in reinforcement learning brain machine interfaces (RMBMIs). First column shows the main author, and the publication year of the reported study is indicated in the next column. Neural decoder type is divided into 3 subcategories including RL base model, function approximator, and learning algorithm. Neural signal and subject types are listed in the consequent columns, along with the number of subjects considered in the RLBMI experiments. Subject column provides gender and specific species if animal study is conducted. Task indicates a type of behavior the external device is required to complete. In addition, the type of external device is listed in the next column. Key reported performance is summarized in terms of success rate, for fair comparisons of all reported studies, and the data amount for evaluation is listed to provide understanding of the learning speed. It should be noted that all provided information from the published studies are summarized. However, there are some missing information as it is not available in the corresponding published studies.

Supplementary Table 1. A summary of reviewed neural decoders in RMBMIs.

Authors	Year	Neural Decoder			Neural Signal Type	Subject	Subject Number	Task	External Device	Closed/ Open Loop	Key Reported Performance	
		RL Base Model	Function Approximator	Learning Algorithm							Best Reported Success Rates	Data Amount for Evaluation
J. Bae, et al.	2011	Q-learning	Kernel Expansion	Kernel Temporal Difference (KTD) (λ)	Intracortical M1 (185 units)	Female Bonnet Macaque	1	2-target center-out reaching task	2D screen	Open	Around 100% after 2 epochs	43 trials/ epoch Average over 50 Monte Carlo runs
J. Bae, et al.	2015	Q-learning	Kernel Expansion	Kernel Temporal Difference (KTD) (λ)	Intracortical M1 (185 units)	Female Bonnet Macaque	1	2-target single step center-out reaching task	2D screen	Open	2-target: 99% after 3 epochs	Average over 50 Monte Carlo runs 43 trials/epoch
								4-target single step center-out reaching task	2D screen	Open	4-target: 99% after 5 epochs	Average over 50 Monte Carlo runs
								8-target single step center-out reaching task	2D screen	Open	8-target: 98% after 6 epochs	Average over 50 Monte Carlo runs 8-target: 178 trials/epoch
								3 target 4 step center-out reaching task	2D screen	Open	Above 60% after 1 epoch	

								4 target 2 step center-out reaching task	2D screen	Open	Above 60% after 1 epoch	
					Intracortical M1 (14 units)	Marmoset Monkey	1	2-target reaching task	Robotic arm	Closed	90% for Day1	20 trials (10 trials each per target)
J. Bae, et al.	2014	Q-learning	Kernel Expansion	Correntropy Kernel Temporal Differences (CKTD)	Intracortical M1 (49 units)	Female Bonnet Macaque	1	4-target center-out reaching task	2D screen	Open	100% after 5 epochs	Average over 50 Monte Carlo runs 144 trials trials/epochs
X. Zhang, et al.	2019	Q-learning	Convolutional Neural Networks	Dueling Deep Q Networks	EEG	Healthy human	7	6 imagery action classification	N/A	Open	Average classification accuracy of 93.63%	34,560 samples per subject
J. DiGiovanna, et al	2007	Watkin's Q(λ)	Feedforward Neural Network	Recursive Least Square	Intracortical	Rat	1	Go no-go task	Robotic arm	Closed	93.7%	1 session 16 trials
J. DiGiovanna, et al.	2007	Watkin's Q(λ)	Feedforward Neural Network	Back-propagation	Intracortical M1 (25 units in left & 33 units in right hemisphere)	Rat	1	2-target reaching task	Robotic arm	Open	max observed 81.3% Avg: 68.1 ± 10.8 (std)%	10 sessions (16 trials/session)
J. DiGiovanna, et al.	2009	Watkin's Q(λ)	Feedforward Neural Network	Back-propagation	Intracortical M1	Male Sprague-	3	2-target reaching task	Robotic arm	Closed	Avg performance: rat01: 68%,	Avg

					(rat01: 16 units, rat02: 17 units, rat03: 29units)	Dawley rat					rat02: 74%, rat93: 73%	2.1±1.2 (std) session (1 session/day) Avg 141.6±41.3(std) trials
J. C. Sanchez, et al.	2011	Watkin's Q(λ)	Feedforward Neural Network	Back-propagation	Intracortical M1 PMd (total 190~240 units)	Female Bonnet Macaque	1	8-target center-out reaching task	2D screen	Open	Reached 100% after 18 epochs	43 trials/epoch
A. Tarigoppula, et al.	2012	Watkin's Q(λ)	Feedforward Neural Network	Back-propagation	Simulated neurons	N/A	N/A	8-target center-out reaching task	2D screen	Open	Over 95% with optimal Izhikevich-tuning depth	80 neurons
Y. Wang, et al.	2015	Attention-Gated Reinforcement Learning (AGREL)	Feedforward Neural Network	Attention-Gated Reinforcement Learning (AGREL)	Intracortical M1 (54 active channels)	Male Rhesus Macaque	1	4-target center-out reaching task	2D screen	Open	Average target acquisition rate reached to 90.16%	Day 1, 2, 3, and 6. 40 mins data/ day. No repetition of the data considered.
X. Shen, et al.	2020	Attention-Gated Reinforcement Learning (AGREL)	Feedforward Neural Network	Attention-Gated Reinforcement Learning (AGREL)	Intracortical M1 (16 channel)	Male Sprague Dawley	6	One lever press task	Lever	Open	Average success rate of 87.5%	for 6 subjects over 300 training epochs

					mPFC (16 channel)							multi day recordings
P. Zhang, et al.	2020	Attention-Gated Reinforcement Learning (AGREL)	Feedforward Neural Network	Transfer Learning and Mini-batch based Attention-Gated Reinforcement Learning (TMAGREL)	Intracortical M1 S1 PPC (monkey01: total 480 neurons Monkey02: total 396 neurons)	Male Rhesus Macaque	2	3-target reaching and grasping task	N/A	Open	Approximately 90% for both monkeys	Monkey01: 600 trails Monkey02: 300 trials
H. Li, et al.	2016	Attention-Gated Reinforcement Learning (AGREL)	Feedforward Neural Network	Maximum Correntropy based Attention-Gated Reinforcement Learning	Intracortical Premotor cortex (55 channels)	Rhesus Macaque	1	4-target obstacle avoidance task	2D screen	Open	Average success rate 68.79%	Total 552 trials for 30 Monte Carlo runs
F. Wang, et al.	2017	Attention-Gated Reinforcement Learning (AGREL)	Kernel Expansion	Quantized Attention-Gated Kernel Reinforcement Learning (QAGKRL)	Intracortical M1 (96 channels) PMd (96 channels)	Male Rhesus Macaque	1	4-target obstacle avoidance task	2D screen	Open	Average success rate of 80.83±10.3%	On one type of learning scenario Total 5000 trials

X. Zhang, et al.	2018	Attention-Gated Reinforcement Learning (AGREL)	Kernel Expansion	Clustering based Kernel Reinforcement Learning	Intracortical 4 simulated neurons	N/A	N/A	4-target reaching task	2D screen	Open	99.8±6.6%	20 Monte Carlo runs for 600 epochs
X. Zhang, et al.	2019	Attention-Gated Reinforcement Learning (AGREL)	Kernel Expansion	Clustering based Kernel Reinforcement Learning	Intracortical M1 (26 channels)	Male Macaque	1	4-target reaching task	Robotic Arm	Open	94.3±0.9%	20 Monte Carlo runs After 400 epochs 1000 datapoint/epoch
X. Zhang, et al.	2019	Attention-Gated Reinforcement Learning (AGREL)	Kernel Expansion	Clustering based Kernel Reinforcement Learning with a weight transfer	Intracortical 3 simulated neurons	N/A	N/A	Two lever discriminative task	Lever	Open	Avg approximately 95%	20 Monte Carlo runs 1000 trials for each task
B. Mahoudi and J. C. Sanchez.	2011	Actor-Critic	Feedforward Neural Network	Back-propagation	Simulated neurons	N/A	N/A	4-target reaching task (2D workspace)	Robotic arm	Closed	Reached 98% after less than 200 trials per target	1 session
					Intracortical M1 (20 units) NAcc (23 units)	Male Sprague-Dawley rat	1	2-target reaching task	Robotic arm	Closed	reached 100% after 16 trials	1 session 40 trials
E. A. Pohlmeyer, et al.	2012	Actor-Critic	Feedforward Neural Network	Hebbian Reinforcement Learning	Intracortical M1	Marmoset Monkey (Callithri	1	2-target reaching task	Robotic arm	Closed	Avg 90% for the first 50 trials	8 sessions

					(21 signals) NAcc (18 signals)	x jacchus)						50~60 trials/session
B. Mahmoudi, et al.	2013	Actor-Critic	Feedforward Neural Network	Hebbian Reinforcement Learning	Simulated neurons	N/A	N/A	2-target center-out reaching task	2D screen	Closed	100% after 2 trials for 2 target task	1 session
								4-target center-out reaching task	2D screen	Closed	100% less than 50 additional trials for 4 target task	1 session
					Intracortical M1 (20 signals)	Marmoset Monkey (Callithrix jacchus)	2	Go no-go task	Robotic arm	Open	Over 95% after 20 trials for both monkeys	3 sessions 1 session/ day
E. A. Pohlmeyer, et al.	2014	Actor-Critic	Feedforward Neural Network	Hebbian Reinforcement Learning	Intracortical M1 (monkey01: avg 18.3±3.1 signals monkey02: avg 21.1±0.4 signals)	Marmoset Monkey (Callithrix jacchus)	2	Go no-go task	Robotic arm	Open	Avg 94%: monkey01 Avg 90%: monkey 02	1000 sessions: monkey01 200 sessions: monkey 02
										Closed	Avg 93%: monkey01 Avg 89%: monkey02	4 sessions 1 session/day
N. W. Prins, et al.	2014	Actor-Critic	Feedforward Neural Network	Hebbian Reinforcement Learning	Intracortical M1 (20 signals)	Marmoset Monkey (Callithri	1	Go no-go task	Robotic arm	Open	From 77% to 83% when Critic accuracy is 90%	100 trials/session

						x jacchus)						

M1: primary motor cortex, mPFC: medial prefrontal cortex, NAcc: nucleus accumbens, PMd: primate dorsal premotor cortex, PPC: posterior parietal cortex, and S1: somatosensory cortex.

Note that in Neural Signal Type, when the input state includes both single and multi-unit activities, a term “signal” was used, as the authors used this term in their studies.