Towards generative AI-based fMRI paradigms: reinforcement learning via real-time brain feedback

Giuseppe Gallitto (giuseppe.gallitto@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Robert Englert (robert.englert@uk-essen.de)
Department of Diagnostic and Interventional Radiology and Neuroradiology
University Medicine Essen
Essen, Germany

Balint Kincses (balint.kincses@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Raviteja Kotikalapudi (raviteja.kotikalapudi@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Jialin Li (jialin.li@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Kevin Hoffschlag (kevin.hoffschlag@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Sulin Ali (sulin.ali@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Ulrike Bingel (ulrike.bingel@uk-essen.de)
Department of Neurology
University Medicine Essen
Essen, Germany

Tamas Spisak (tamas.spisak@uk-essen.de)
Department of Diagnostic and Interventional Radiology and Neuroradiology
University Medicine Essen
Essen, Germany

Abstract

Traditional fMRI's reliance on fixed task paradigms perpetuates the reverse inference problem, limiting specificity in brain-behavior mapping. We present Reinforcement Learning via Brain Feedback (RLBF), an approach reversing the direction of inference by using real-time fMRI and reinforcement learning to dynamically adjust stimuli – optimizing the 'stimulus space'. In a visual cortex proof-of-concept study (N=10), the algorithm successfully optimized checkerboard parameters within 35 trials, improving brain prediction from chance-level to a mean absolute percentage error (MAPE) of 12.7% (SD: 6.3%; inter-trial improvement of 0.6%) and achieving stable convergence, despite fMRI noise. Stimulation optimization revealed a preference for maximum contrast stimuli at 18Hz (+/-10Hz across participants) – aligning with known visual processing properties – and demonstrated the potential for brain activity to effectively tune AI models, offering new avenues for personalized experimental design and rigorous testing of reverse inference claims.

Keywords: real-time fmri; reinforcement learning; experimental paradigms; generative ai

Introduction

In traditional human neuroimaging experiments, researchers construct experimental paradigms with a certain psychological/behavioral construct validity to infer the corresponding neural correlates. Here we introduce a novel approach called Reinforcement Learning via Brain Feedback (RLBF), that inverts the direction of inference; it searches for the optimal stimulation parameters to maximize (or minimize) response in predefined brain regions or networks (fig.1). The "stimulus space" is optimized by a reinforcement learning algorithm (Kaelbling et al., 1996) which is rewarded based on real-time fMRI (Sulzer et al., 2013) data. Specifically, during ongoing real-time fMRI acquisition, the reinforcement learning agent manipulates the stimulus space (e.g. by means of generative AI) to drive the participant's neural activity in a specific direction. Then the agent is rewarded based on the measured brain responses and gradually learns to adjust its choices to converge towards an optimal solution. Here, we present the results of a proof of concept study that aimed to confirm the viability of the proposed approach with simulated and empirical real time fMRI data.

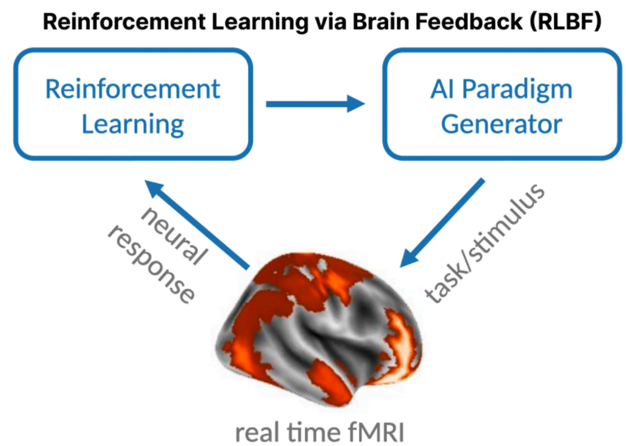

Figure 1. ***Concept behind RLBF.*** *Instead of using an experimental paradigm, we represent a broader paradigm space by generative modelling. This generative model is controlled or fine-tuned by reinforcement learning, resulting in an approximately optimized stimulation strategy.*

Methods

In our proof of concept study, we aimed to construct a streamlined setup.

Reinforcement Learning: To implement the reinforcement learner (fig 1. "Reinforcement Learning"), we used a simple and widely used algorithm, a soft Q-learner (Haarnoja et al., 2017) with a smooth reward function.

Stimulus Space: Participants viewed multiple iterations of a flickering checkerboard, where contrast and frequency (values between 0 and 1) served as free parameters within the stimulus space (Fig. 1, "AI Paradigm Generator"). Contrast value of zero resulted in no difference to the resting block. The reward signal for the reinforcement learner was calculated from brain responses in the primary visual cortex (V1), as measured by a linear model fitted on a single block of data measured in a block-design fashion, with 5 seconds of visual stimulus followed by 11 seconds of rest. The hypothesis function was convolved with the canonical SPM HRF. In this setting, the task for the agent was to figure out the optimal contrast-frequency configuration that maximizes a participant's brain activity in V1.

In-silico Hyperparameter Tuning: we defined the optimal ground truth as a linear function of contrast and flickering frequency, with maximum activation at maximal contrast and an arbitrary frequency value of 0.7. In one simulation run, the reinforcement learner had 100 trials. In each trial the agent picked a contrast and frequency value and updated its Q-table based on the reward that was calculated by our ground truth equation, with added Gaussian noise. We fine-tuned the hyperparameters for the models using realistic parameter initialization (signal-to-noise: 0.5 - 5; q-table smoothing σ: 0.5 - 4.0; soft-Q temperature: 0.2; learning rate: 0.05 - 0.9).

Study Participants: with parameters chosen based on our simulation results, we measured data in n=10 participants (M=5, F=5; mean age=25,9, age SD=3.93), to establish the proof of concept. The first N=3 participants' fMRI runs were also utilized to further improve RL hyperparameters (manually adjusted).

Software: the RLBF methodology is implemented as a python package (Gallitto et al., n.d.-a).

Results and Discussion

Simulation results show that the proposed implementation provides robust solutions in a relatively wide range of initial conditions, within a small number of trials (see Table 1). Overall, high q-table smoothing (σ=4.0) appears to function well with SNRs >= 2.0, with lower learning rates (0.02) and temperature (0.08) for optimal training. The model displayed a remarkable stability with decreasing prediction error (MAPE) across trials (Table 1). Results from the empirical measurements are in line with knowledge about the contrast and frequency dependence of the checkerboard response (Albrecht & Hamilton, 1982; Victor et al., 1997, Albrecht et al., 2003) and provide initial confirmation for the feasibility of the proposed approach (code, analyses and empirical data are disclosed in our supporting repository, see Gallitto et al., n.d.-b).

We've introduced Reinforcement Learning with Brain Feedback (RLBF), a novel experimental approach to find optimal brain stimulation for modulating individual brain activity. This proof-of-concept used a simplified setup, but future work will focus on extending it to generative AI solutions. By inverting inference from ("brain -> behavior"; instead of "behavior -> brain") RLBF could become a new tool for basic and translational research. When paired with generative AI, RLBF has the potential to offer novel individualized treatments, such as AI-generated text, video, or music optimized for e.g. improving mental states like e.g. pain or anxiety.

Subject	Run	MAPE	SNR	Q-table Max
sub-001	1	-0.0165	3.75	[0.9, 0.5]
sub-001	2	-0.0090	3.39	[0.9, 0.0]
sub-002	1	-0.0015	2.85	[0.9, 0.9]
sub-002	2	-0.0013	3.48	[0.9, 0.0]
sub-003	1	-0.0138	2.91	[0.9, 0.0]
sub-003	2	-0.0072	4.63	[0.9, 0.9]
sub-004	1	-0.0086	6.85	[0.9, 0.0]
sub-004	2	-0.0095	4.94	[0.9, 0.9]
sub-005	1	-0.0157	3.18	[0.9, 0.9]
sub-005	2	-0.0050	4.15	[0.9, 0.9]
sub-006	1	-0.0065	4.03	[0.9, 0.0]
sub-006	2	-0.0169	4.28	[0.9, 0.9]
sub-007	1	-0.0142	3.25	[0.9, 0.0]
sub-007	2	0.0085	3.13	[0.9, 0.9]
sub-008	1	-0.0014	4.02	[0.9, 0.9]
sub-008	2	-0.0037	4.09	[0.9, 0.9]
sub-009	1	0.0026	4.52	[0.9, 0.9]
sub-009	2	-0.0071	4.06	[0.9, 0.9]
sub-010	1	-0.0039	3.72	[0.9, 0.0]
sub-010	2	0.0002	4.20	[0.9, 0.0]

Table 1: Regardless of data noisiness, the algorithm consistently achieved maximum contrast at the 35th trial (see Q-table Max as [contrast, frequency]), showing moderate improvement slopes (MAPE).

Acknowledgments

The work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — Project-ID 422744262 – TRR 289 (Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) — Projektnummer 422744262–TRR 289) and Projektnummer 316803389 – SFB 1280 "Extinction Learning".

References

Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2003). Nonlinear properties of visual cortex neurons: Temporal dynamics, stimulus selectivity, neural performance. *The visual neurosciences*, *1*, 747–764.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*, *48*(1), 217–237.

Gallitto, G., Englert, R., Kincses, B., Kotikalapudi, R., Li, J., Hoffschlag, K., Ali, S., Bingel, U., & Spisak, T. (n.d.-a). *Reinforcement Learning via Brain Feedback (RLBF)* [Software]. Retrieved from https://github.com/pni-lab/RLBF

Gallitto, G., Englert, R., Kincses, B., Kotikalapudi, R., Li, J., Hoffschlag, K., Ali, S., Bingel, U., & Spisak, T. (n.d.-b). *Supporting repository for the manuscript "Towards generative AI-based fMRI paradigms: reinforcement learning via real-time brain feedback"* [Software]. Retrieved from https://github.com/gallg/RLBF_analysis

Haarnoja, T., Tang, H., Abbeel, P., & Levine, S. (2017, July). Reinforcement learning with deep energy-based policies. In *International conference on machine learning* (pp. 1352–1361).

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, *4*, 237–285.

Sulzer, J., Haller, S., Scharnowski, F., Weiskopf, N., Birbaumer, N., Blefari, M. L., et al. (2013). Real-time fMRI neurofeedback: progress and challenges. *NeuroImage*, *76*, 386–399.

Victor, J. D., Conte, M. M., & Purpura, K. P. (1997, May). Dynamic shifts of the contrast-response function. *Visual Neuroscience*, *14*(3), 577–587.