bioRNNs: Toward More Biologically Plausible Neural Network Models of the Brain

Ahmad Beyh[1], Jason Z. Kim[2], Linden Parkes[1]

[1]Department of Psychiatry, Brain Health Institute, Rutgers University, Piscataway, NJ | [2]Department of Physics, Cornell University, Ithaca, NY

Contact: a.beyh@rutgers.edu

Background

Understanding how brain structure shapes neural dynamics is fundamental to explaining cognition, development, and disease.

Recurrent neural networks (RNNs) are powerful tools for modeling neural dynamics, but **typical architectures exhibit the following critical departures from biology**:

- **All-to-all connectivity:** Every node processes inputs and contributes to outputs
- **No spatial constraints:** Hidden layer connections ignore brain organization

Objectives

We developed **biophysical recurrent neural networks (bioRNNs)** that address these limitations through **two key constraints**:

1. **I/O Masking:** Restricts inputs to sensory cortex and outputs to association cortex
2. **Spatial Embedding:** Constrains connectivity based on the sensorimotor-association gradient

Methods Overview

Network Architecture

- **100 nodes** representing cortical regions (Schaefer atlas)
- Embedding uses **inter-regional distances along the cortical hierarchy**

Training

- 100 bioRNNs trained on **Go/No-Go** and **Delayed Match to Sample** tasks
- Compared against physical embedding and unconstrained RNNs

Validation

- Intrinsic neural timescales compared to empirical resting-state fMRI from HCP data

Network Architecture

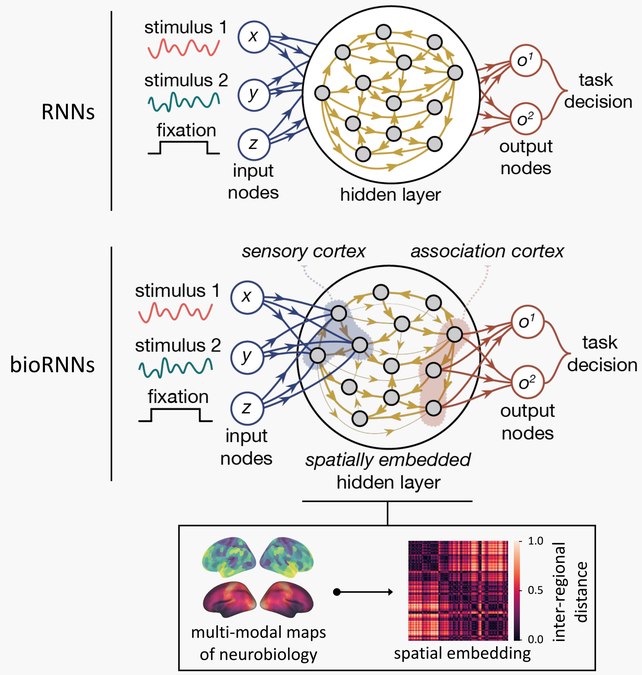

Figure 1. (Top) Classical RNNs employ all-to-all connectivity with unconstrained hidden layers. (Bottom) Biophysical RNNs implement I/O masking and spatial embedding based on multi-modal neurobiological maps.

Training Process

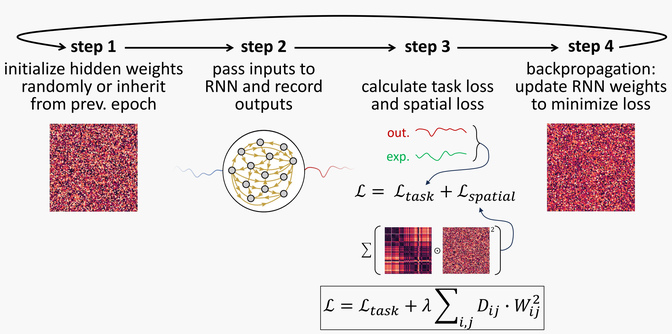

Figure 2. Training incorporates a spatial loss term that penalizes connections between distant nodes, encouraging biologically plausible weight structures.

Results

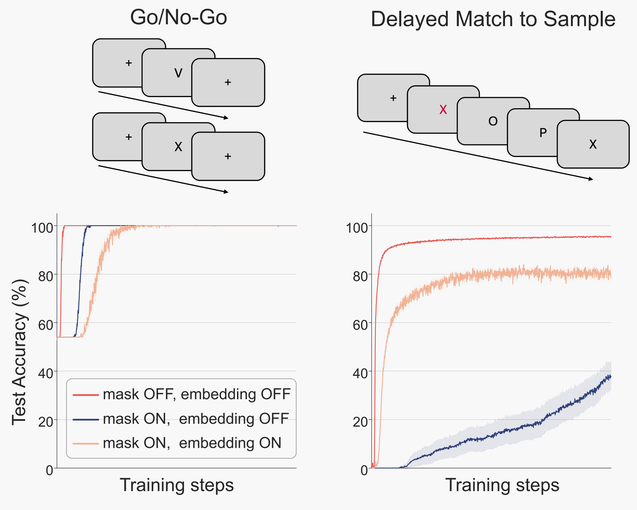

Figure 3. Training performance on Go/No-Go (left) and Delayed Match to Sample (right). For simple tasks, masking alone suffices. For complex tasks, **embedding is crucial for recovering performance** when masking is applied.

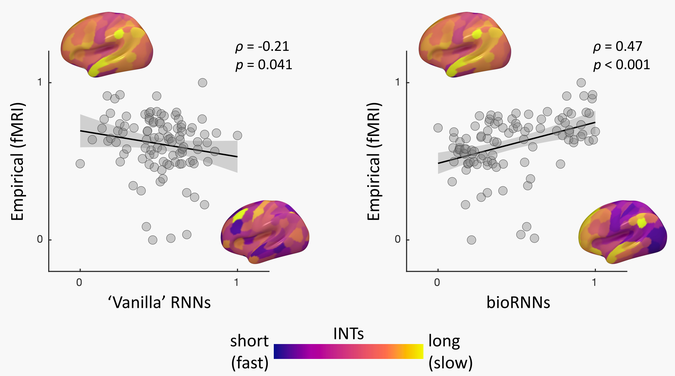

Figure 4. Simulated vs. empirical intrinsic neural timescales (INTs). Vanilla RNNs show weak negative correlation (rho = -0.21), while **bioRNNs exhibit strong positive correlation (rho = 0.47, p less than 0.001)**, indicating biologically realistic dynamics.

Key Findings

- **Biological constraints provide useful inductive biases** that accelerate learning
- **Embedding rescues performance** for complex tasks when I/O masking is applied
- **bioRNNs exhibit neural dynamics that closely match empirical fMRI** (r = 0.47, p less than 0.001)

Conclusions

Incorporating simple biological constraints into RNNs allows them to capture fundamental aspects of neural computation.

bioRNNs provide versatile experimental platforms bridging neuroscience discovery and AI advances.

Future work: Incorporate more nuanced biological constraints and interrogate structure-function relationships in psychiatric conditions.

Implications

- **Systems neuroscience:** Tools for testing how brain structure constrains function
- **Developmental neuroscience:** Modeling cortical hierarchy emergence
- **Clinical applications:** Disease modeling by perturbing biological constraints

References

1. Kim JZ, et al. (2024) Biochem Biophys Res Commun 728:150302
2. Mante V, et al. (2013) Nature 503:78-84
3. Achterberg J, et al. (2023) Nat Mach Intell 5:1369-1381
4. Wang PY, et al. (2021) Neuron 109:3879-3892.e5
5. Yang GR, et al. (2019) Nat Neurosci 22:297-306
6. Schaefer A, et al. (2018) Cereb Cortex 28:3095-3114
7. Sydnor VJ, et al. (2021) Neuron 109:2820-2846

Acknowledgments: This work was supported by the Brain Health Institute Travel Award and the National Institute of Mental Health (NIMH).