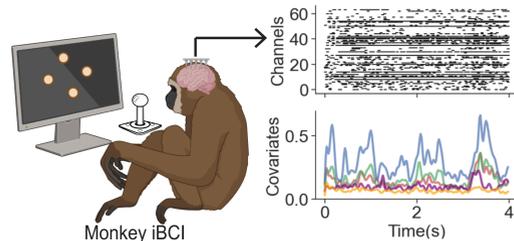# SPINT: Spatial Permutation-Invariant Neural Transformer for Consistent Intracortical Motor Decoding

T. Le[1], H. Fang[1], J. Li[1], Tung Nguyen[2], Lu Mi[3], A. L. Orsborn[1], U. Sümbül[4], E. Shlizerman[1]
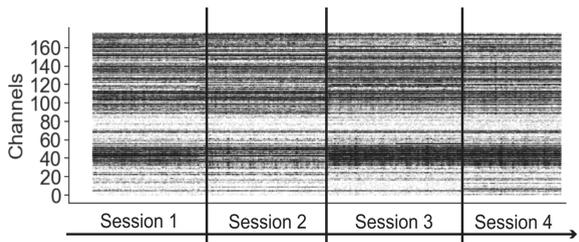
[1]University of Washington, Seattle, WA, USA. [2]University of California, Los Angeles, CA, USA. [3]Georgia Institute of Technology, GA, USA. [4]Allen Institute, Seattle, WA, USA
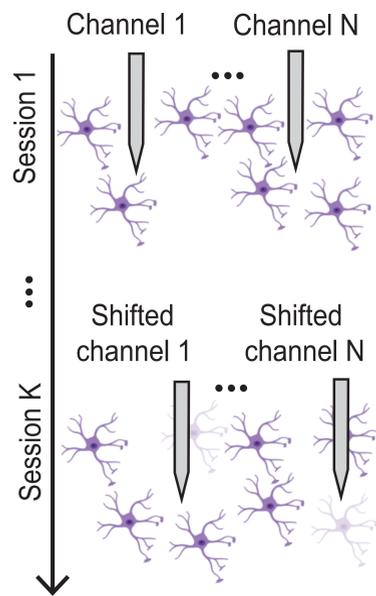
Introductions

- BCIs rely on algorithms for inference from neural siganls, often called *decoders*.

Monkey iBCI

- Long-term deployment of iBCIs suffers from the nonstationarity of neural recording.

Nonstationary neural activity over sessions

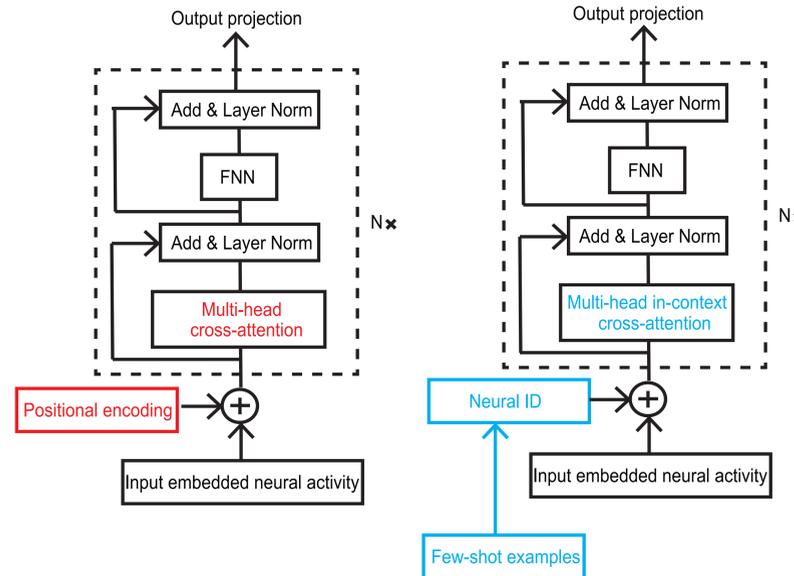- Current algorithms relies on explicit alignment techniques, which

  A. require model finetuning during test time.
  B. have limlited generalizability across sessions.
  C. imporse heavy computational burden

Goal:
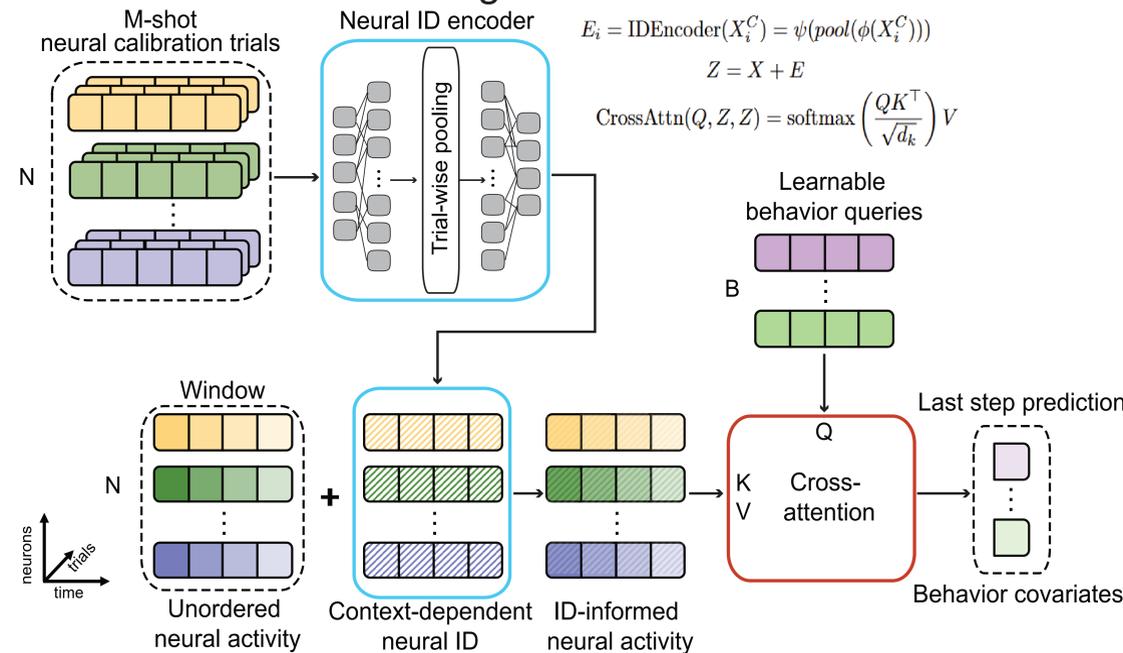
Develop a novel transformer model that can adapt to different sessions during test time without finetuning any model parameters.

Methods

- From standard positional embedding to neural ID embedding

- Neural transformer deisgn

$$E_i = \text{IDEncoder}(X_i^C) = \psi(pool(\phi(X_i^C))) \quad (1)$$

$$Z = X + E \quad (2)$$

$$\text{CrossAttn}(Q, Z, Z) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V \quad (3)$$

Unordered neural activity · Context-dependent neural ID · ID-informed neural activity · Cross-attention · Last step prediction · Behavior covariates

Learnable behavior queries
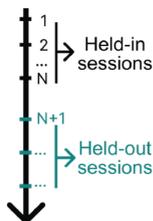
- Permutation invariant property

Proposition 1. *Cross-attention with identity-informed neural activity (Equation 3) is invariant to the permutation of neural units, i.e.,*

$$\text{CrossAttn}(Q, Z, Z) = \text{CrossAttn}(Q, P_R Z, P_R Z), \quad (4)$$

where P_R is the row permutation matrix. (See proof in Appendix).

- Evaluation pipelines

  1. Within-session
  2. Cross-session

  Held-in sessions / Held-out sessions

- Methods

  1. Oracle (OR)
  2. Zero-shot (ZS)
  3. Few-shot supervised (FSS)
  4. Few-shot unsupervised (FSU)
  5. Ours

Experimental results

- SOTA in within-session evaluation

	Class	M1	M2	H1
Wiener Filter (WF)	OR	0.54 ± 0.01	0.27 ± 0.02	0.24 ± 0.02
RNN	OR	0.75 ± 0.03	0.59 ± 0.07	0.51 ± 0.09
NDT2 Multi [1]	OR	0.77 ± 0.03	0.62 ± 0.03	0.68 ± 0.05
NDT2 Multi [1]	FSS	0.77 ± 0.03	0.63 ± 0.03	0.62 ± 0.04
WF	ZS	0.46 ± 0.06	0.15 ± 0.07	0.20 ± 0.04
RNN	ZS	0.52 ± 0.15	0.20 ± 0.29	0.31 ± 0.13
CycleGAN + WF [2]	FSU	0.61 ± 0.02	0.32 ± 0.03	0.15 ± 0.04
NoMAD + WF [3]	FSU	0.64 ± 0.01	0.35 ± 0.05	0.21 ± 0.06
SPINT (Ours)	GF-FSU	$\mathbf{0.77 \pm 0.02}$	$\mathbf{0.59 \pm 0.01}$	$\mathbf{0.47 \pm 0.06}$

- SOTA in cross-session evaluation

	Class	M1	M2	H1
Wiener Filter (WF)	OR	0.53 ± 0.04	0.26 ± 0.03	0.21 ± 0.04
RNN	OR	0.75 ± 0.05	0.56 ± 0.04	0.44 ± 0.13
NDT2 Multi [14]	OR	0.78 ± 0.04	0.58 ± 0.04	0.63 ± 0.08
NDT2 Multi [14]	FSS	0.59 ± 0.07	0.43 ± 0.08	0.52 ± 0.04
WF	ZS	0.34 ± 0.06	0.06 ± 0.04	0.16 ± 0.03
RNN	ZS	-0.60 ± 0.45	-0.07 ± 0.23	0.09 ± 0.18
CycleGAN + WF [23]	FSU	0.43 ± 0.04	0.22 ± 0.06	0.12 ± 0.06
NoMAD + WF [19]	FSU	0.49 ± 0.03	0.20 ± 0.10	0.13 ± 0.10
SPINT (Ours)	GF-FSU	$\mathbf{0.66 \pm 0.07}$	$\mathbf{0.26 \pm 0.13}$	$\mathbf{0.29 \pm 0.15}$

- Ablation study for neural ID embedding

	M1	M2	H1
Without PE	0.50 ± 0.02	-0.09 ± 0.09	0.09 ± 0.06
Absolute PE	0.46 ± 0.01	-0.06 ± 0.09	0.09 ± 0.09
Context-dependent ID	$\mathbf{0.66 \pm 0.07}$	$\mathbf{0.26 \pm 0.13}$	$\mathbf{0.29 \pm 0.15}$

- Investigation of number of trials (A) and training days (B)

Conclusions and future directions

- Toward future plug-and-play robust iBCI decoder.

- Supports causal and real-time decoding (in-silico simulation) but is needed to be evaluated in vivo.

- Extend to more complicated decoding task, e.g., communication or speech

Acknowledgement