 DOI: 10.3217/978-3-85125-962-9-170

BRAND: A platform for real-time deep network inference in closed-loop BCIs

YH Ali[1,2], K Bodkin[3], M Rigotti-Thompson[1,2], K Patel[5], B Bhaduri[2], SR Nason-Tomaszewski[1], DM Mifsud[1], X Hou[5], C Nicolas[6], S Allcroft[7], LR Hochberg[6,7,8,9], N Au Yong[1], SD Stavisky[5], LE Miller[3,4], DM Brandman[†5], C Pandarinath[*†1,2]

[1]*Emory Univ.* [2]*Georgia Inst. of Tech., Atlanta, GA, USA;* [3]*Northwestern Univ.* [4]*Shirley Ryan AbilityLab, Chicago, IL, USA;* [5]*Univ. of California, Davis, CA, USA;* [6]*Mass. General Hospital, MA, USA;* [7]*Brown Univ., RI, USA;* [8]*Harvard Med. School, MA, USA;* [9]*Dept. of Veterans Affairs Med. Center, RI, USA*

[†]Equal Contribution. *1760 Haygood Dr NE, Atlanta, GA, USA. E-mail: chethan [at] gatech.edu.

Introduction: Closed-loop experiments are key components of brain-computer interface (BCI) research. Artificial neural networks (ANN) are state-of-the-art tools for modeling and decoding neural activity, but deploying them in closed-loop experiments is challenging. Researchers need a framework that supports high-level programming languages for running ANN (e.g., Python and Julia) while maintaining support for languages critical for low-latency data acquisition and processing (e.g., C and C++). To address these needs, we introduce the BRAND Realtime Asynchronous Neural Decoding system (BRAND).

Materials, Methods, and Results: BRAND can run on almost any standard Linux computer and comprises processes, termed *nodes*, that communicate with each other via streams of data in a *graph*. BRAND supports reliable real-time execution with microsecond precision, making it an ideal platform for closed-loop neuroscience and neural engineering applications. BRAND uses Redis [1] to send data between nodes, which enables fast inter-process communication (IPC), support for 54 programming languages, and distributed processing across multiple computers. Developers can deploy existing ANN models seamlessly in BRAND with minimal implementation changes. In initial testing, BRAND achieves a fast IPC latency (<500 microseconds) when sending large quantities of data (1024 channels of 30 kHz simulated neural data in 1 ms blocks). BCI control was tested with a graph that receives 30 kHz microelectrode array voltage recordings via Ethernet, filters and thresholds the input to get spikes, bins spikes into 10 ms bins, applies a decoding model, and updates the position of a cursor on a display. In an initial demonstration of the system, participant T11 in the BrainGate2 clinical trial (NCT00912041) achieved a target acquisition time of 2.84 ± 0.83 seconds (53 trials) on a radial-8 center-out cursor control task, in which 30 kHz signal processing, linear decoding, task control, and graphics were all executed in BRAND. Future experiments will incorporate ANN; to benchmark ANN latency, we ran a PyTorch-based recurrent neural network decoder (10 hidden units, 30-bin input sequences) and measured latency (N = 30,000 packets). The end-to-end latency from signal input to decoder prediction was consistently less than 2 ms for this configuration (**Fig. 1**). We also validated that BRAND can run two popular neural population dynamics models – Latent Factor Analysis via Dynamical Systems (LFADS) [2] and Neural Data Transformer (NDT) [3] – in real-time, with latencies below 6 ms per 10 ms bin (256-channel data), using their original Tensorflow and PyTorch implementations.

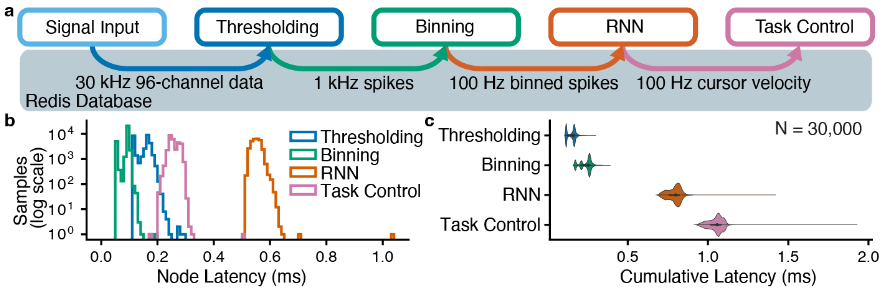

Figure 1*. a) BRAND nodes in a graph. b) Latency of each node. c) End-to-end latency of the system.*

Discussion: BRAND supports low-latency ANN inference while providing seamless integration with the data acquisition, signal processing, and task code that is needed for closed-loop BCI research.

Significance: With its modular design and broad language support, BRAND simplifies the process of translating computational models from offline analysis into closed-loop experiments that leverage the power of ANNs to improve BCI control across several contexts.

Acknowledgements: This work was supported by the Emory Neuromodulation and Technology Innovation Center (ENTICe), NSF NCS 1835364, DARPA PA-18-02-04-INI-FP-021, NIH Eunice Kennedy Shriver NICHD K12HD073945, NIH-NINDS/OD DP2NS127291, the Alfred P. Sloan Foundation, the Burroughs Wellcome Fund, the Simons Foundation as part of the Simons-Emory International Consortium on Motor Control (CP), NIH NINDS NS053603, NS074044 (LEM), NIH NIBIB T32EB025816 (YHA), NIH-NIDCD U01DC017844, and the Department of Veterans Affairs Rehabilitation Research and Development Service A2295R (LRH).

References: [1] Redis https://redis.io/. [2] Pandarinath et al. 2018, Nat Methods doi:10.1038/s41592-018-0109-9. [3] Ye and Pandarinath 2021, Neurons Behav Data Analysis Theory doi:10.1101/2021.01.16.42695.

 Article ID: 144565