DOI: 10.3217/978-3-99161-050-2-103

Efficient and accurate cortical spike train decoding for BCI implants with recurrent spiking neural networks

Tengjun Liu[1,2], Tuoru Li[1], Julia Gygax[2], Julian Rossbroich[2], Tianyang Li[3], Yuqi Yang[1], Zihao Li[1], Yansong Chua[4*], Jie Yang[3], Weidong Chen[1], Friedemann Zenke[2], Shaomin Zhang[1]

[1]*Zhejiang University, Hangzhou, China;* [2]*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland;* [3]*Westelake University, Hangzhou, China;* [4]*China Nanhu Academy of Electronics and Information Technology, Jiaxing, China*

*111 Qixing Road, Jiaxing, China. E-mail: caiyansong@cnaeit.com

Introduction: Externally mounted pedestals of invasive brain-computer interfaces (BCIs) risk infection, requiring fully implanted systems. These systems must meet strict latency and energy limits with reliable decoding. While recurrent spiking neural networks (RSNNs) are well-suited for low-power neuromorphic hardware, it is unclear if they can provide both high decoding performance and low energy use.

Material, Methods and Results: To evaluate this, we trained a tinyRSNN to decode finger velocity from cortical spike trains (CSTs) of two monkeys (Fig. 1A-C). Our tinyRSNN outperformed classical Kalman Filter (KF) and max coefficient of determination (R^2) baselines in NeuroBench [1] (ANN3d, SNN3), though it fell behind AEGRU, the top R2 achiever in the IEEE BioCAS 2024 neural decoding challenge [2]. However, we can achieve comparable R^2 by scaling up our model to bigRSNN (Fig. 1D). The tinyRSNN, notably, consumed less energy in synaptic operations than the baselines for balancing high R^2 with low energy need in NeuroBench (ANN2d, SNN2) [1], while achieving much higher R^2. In particular, tinyRSNN consumed only around $1/2557$ AEGRU's effective energy (Fig. 1E). To determine what contributes to this good trade-off, we show in ablation studies that dynamic synapses, recurrent connections, learnable heterogeneous time constants, and pretraining improved R^2, while synaptic pruning and activity regularization reduced energy use (Fig. 1F). Finally, tinyRSNN was deployed on an FPGA (VCU118), achieving low power (around 12mW) and similar R^2 to GPU-based results (Fig. 1G).

Conclusion: Our results thus show that tinyRSNN offers competitive CST decoding performance under tight resource constraints and can be deployed on low-power FPGAs without a loss in decoding accuracy.

Acknowledgments and Disclosures: This work is funded by: ZJU Doctoral Graduate Academic Rising Star Development Program (No. 2023059), EU's Horizon Europe Research and Innovation Programme (Grant Agreement No. 101070374, CONVOLVE) funded through SERI (Ref. 1131-52302), the Swiss National Science Foundation (Grant Number PCEFP3_202981), and the Novartis Research Foundation.

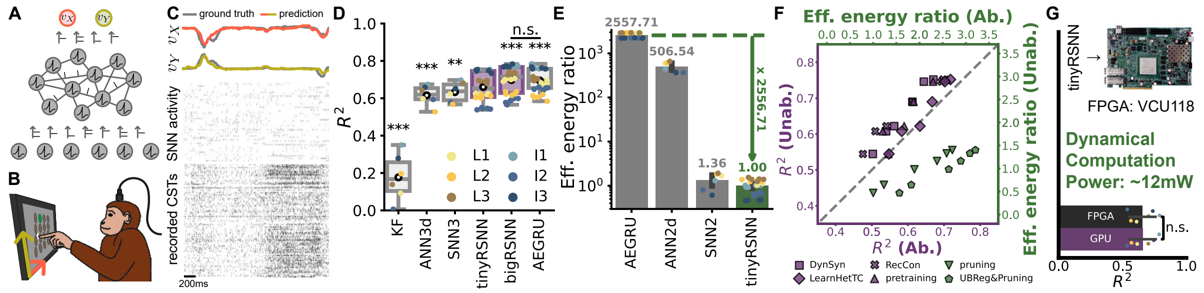

Figure 1: ***A****. RSNN model: one input layer, one recurrent hidden layer (leaky integrate-and-fire neurons), and one output layer (leaky integrator neurons).* ***B****. Two monkeys (L & I) perform a random dot reaching task [3].* ***C****. RSNN activities: Bottom: recorded CSTs (input layer); middle: hidden layer spike raster; Top: output layer membrane potentials (predictions) vs. ground truth.* ***D****. Decoding performance (R^2) of RSNN models vs. baselines. * denotes statistical significance of the proposed tinyRSNN and other models, based on paired t-test: **: $p < 0.01$, ***: $p < 0.001$. n.s.: not significant. Results are shown from six sessions (L1-L3, I1-I3) and five random initializations each.* ***E****. Comparison of energy consumption (effective energy ratio).* ***F****. Ablation studies for decoding accuracy (purple) and energy consumption (green). Each point (representing one session) should be interpreted according to the axis of its corresponding color. Ab. & Unab.: Ablated & Unablated results. DynSyn: dynamical synapses, LearnHetTC: learnable heterogeneous time constants, RecCon: recurrent connections, UBReg: upper bound neuronal activity regularization.* ***G****. tinyRSNN implementation on the FPGA (VCU118) with its decoding accuracy and energy consumption.*

References:

[1] Yik, J., Frenkel, C. & Reddi, V. J. Advancing Neuromorphic Computing Algorithms and Systems with NeuroBench. In *NeurIPS 2024 Workshop Machine Learning with new Compute Paradigms*, 2024.

[2] Zhou, B., Sun, P. S. V., Yik, J., et al. Grand Challenge on Neural Decoding for Motor Control of non-Human Primates. In *2024 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2024.

[3] O'Doherty, J. E., Cardoso, M. M. B., Makin, J. G., et al. Nonhuman Primate Reaching with Multichannel Sensorimotor Cortex Electrophysiology [Data set]. *Zenodo*, 2017.

Published by Verlag der Technischen Universität Graz