

Sustainable AI

Energy costs and alternatives
to brute-force scaling

Prof. Dr. Sebastian Otte

Adaptive AI Lab
University of Lübeck

CORE MESSAGE

Scaling has a resource bill

1,764

kWh per week

Illustrative (small) GPT-2 training setup
8 GPUs, 7 kW server load, 1.5x overhead

**More parameters push compute,
hardware demand and electricity
use upward together.**

The talk frames efficiency as a learning-system
problem: better dynamics can reduce
dependence on brute-force scale.

7 kW x 1.5 x 24 h x 7 days

A worked example from the talk, not a universal training benchmark

Learning from biological computation

The research direction places temporal dynamics and hardware behavior at the center of learning.

01

Spikes

Communicate events

02

Resonance

Select temporal structure

03

Recurrence

Maintain state

04

Device physics

Compute through material dynamics

PAPER 01 / ICML 2024

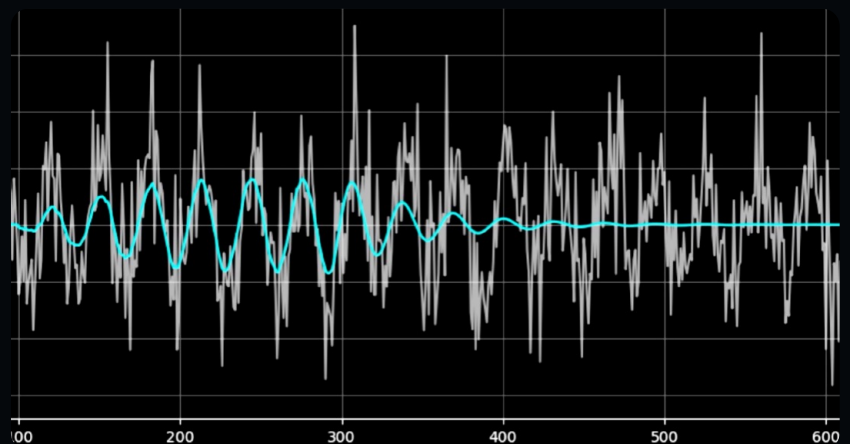
Balanced Resonate-and-Fire Neurons

A resonator responds strongly to matching temporal patterns while suppressing off-pattern noise.

Balancing keeps the oscillatory state centered after spikes, giving the neuron a stable operating regime.

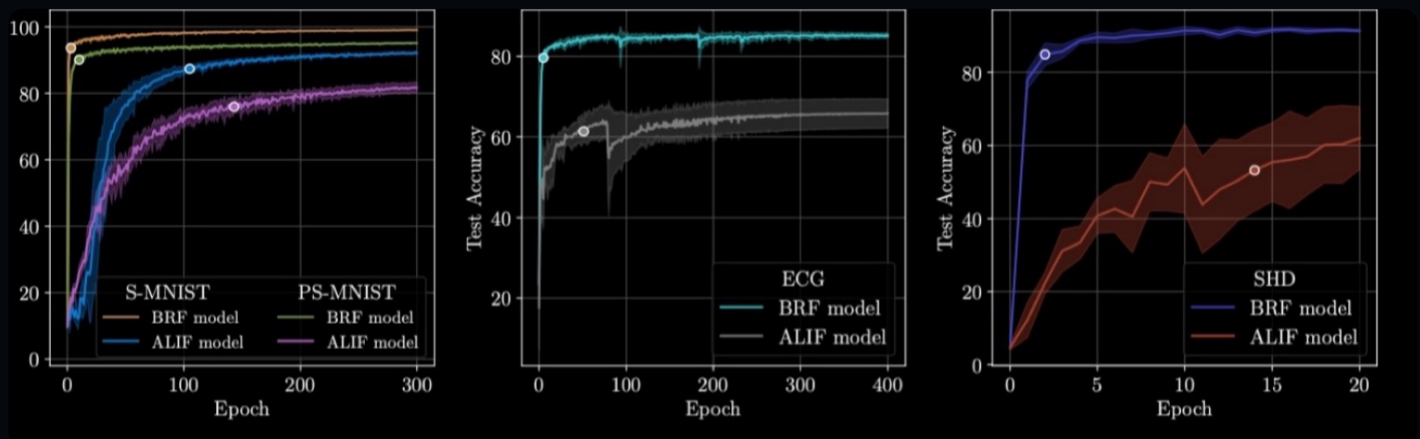
Noisy input

Resonator state



The cyan trace isolates a short resonant episode from a long noisy sequence.

BRF performance and application



99.0 ± 0.1%

S-MNIST test accuracy

0.22x

SOP ratio versus ALIF on S-MNIST

Example: WindSenseAI

Vibration-based anomaly detection for wind turbines

S. Higuchi, S. Kairat, S. Bohté, S. Otte, "Balanced Resonate-and-Fire Neurons," ICML 2024.

Adaptive AI Lab

05

MARS: dynamics shaped for memristors

Inputs pass through an encoder and stacked memristor blocks.

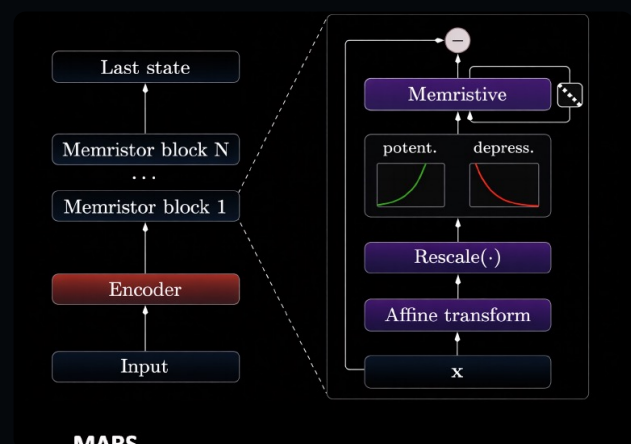
Potentiation and depression curves define the state update. The formulation exposes a parallelizable recurrence for time-series classification.

Memristive-friendly

Uses device behavior as part of the computation

Parallelizable

Designed for scalable sequence processing



C. C. Horuz, A. Ceni, C. Gallicchio, S. Otte, "Scalable Memristive-Friendly Reservoir Computing for Time Series Classification," arXiv:2604.19343, 2026.

Adaptive AI Lab

06

MARS cuts model size and training time

2,580x

fewer parameters

401 MARS
1,034,242 Mamba

3,025x

shorter training time

0.2 s MARS
605 s Mamba

Across the six datasets shown in the talk

MARS used 401-1,604 parameters and trained in 0.2 to 1.6 seconds.

Ratios are derived from the highlighted Heartbeat row. Hardware and implementation context remain part of the full paper.

C. C. Horuz, A. Ceni, C. Gallicchio, S. Otte, "Scalable Memristive-Friendly Reservoir Computing for Time Series Classification," arXiv:2604.19343, 2026.

Adaptive AI Lab

07

GROUP MISSION

Adaptive AI Lab mission

Mission wording inferred from the talk

Develop learning systems that achieve more with less computation by combining biological inspiration with device-aware design.

Scientific aim

Understand how temporal dynamics, memory and sparsity shape learning.

Engineering aim

Translate those principles into reliable, efficient systems for real signals.

Research directions

Directions inferred from the examples in the talk

- | | | |
|----|------------------------------------|--|
| 01 | Efficient temporal learning | Sequence and time-series models that exploit temporal structure. |
| 02 | Neuromorphic computation | Spiking and resonant dynamics for sparse event processing. |
| 03 | Physical intelligence | Algorithms aligned with the behavior of emerging devices. |
| 04 | Energy systems | Monitoring and earlier fault detection. |

CONTACT

Prof Dr. Sebastian Otte

Adaptive AI Lab
Institute for Robotics and Cognitive Systems
University of Lübeck

sebastian.otte@uni-luebeck.de

[linkedin.com/in/sebastian-otte](https://www.linkedin.com/in/sebastian-otte)

Featured work

Balanced Resonate-and-Fire Neurons, ICML 2024
MARS, arXiv:2604.19343, 2026

