UNIVERSITÄT DUISBURG ESSEN

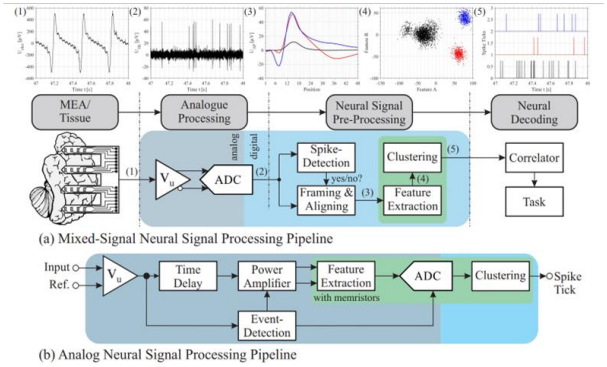
Offen im Denken

Theorie-Kolloquium SS 2026
Fr 29.05.2026, 14:00-15:30
MC 351 & online (URL in E-Mail)

udue.de/TheoKoll

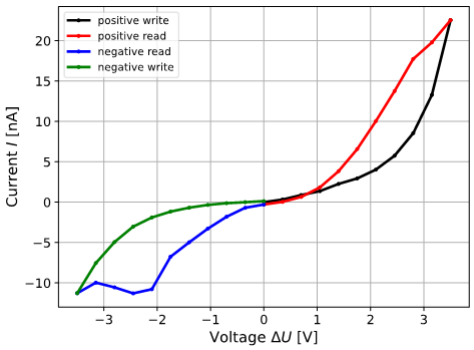
Can tiny deep learning hardware solve real-world problems?

Dr.-Ing. Andreas Erbslöh, UDE



Left: Example of two signal processing pipeline for reading and processing neural activities from a tissue (from sensor input to control output) – (a) classical analogue pre-processing with digital preprocessing and AI inference – (b) memristor-based computing

Right: Current-Voltage Relationship of a memristor with identification of read and write branches for defining specific state



Current artificial intelligence (AI) systems have made significant progress in recent years. Deep neural networks (DNNs) and transformer-based architectures have been used in modern large language models / generative AI tools. Despite their capabilities, these approaches have fundamental structural limitations: such complex models are executed in large data centres with hundreds of GPU nodes. For example, the training of ChatGPT-4 (1.8 trillion parameters in 120 layers, combining 16 different task-specific LLMs) cost 80 million USD and consumed 62.32 GWh over 100 days [1]. By comparison, households in California had an energy consumption of 2 MWh over the same period [2]. Overall, these models are highly data- and resource-intensive, places high demands on IT infrastructure, the models are difficult to interpret, and they suffer from a lack of robustness and exhibit unpredictable behaviour when input data is affected by drift. For application-specific use in resource-constrained systems for local sensor data processing, the entire chain from data acquisition, data processing and data annotation to data flow design and the design of AI models must be conceived and optimised as an end-to-end AI system. To solve this, two major challenges are available:

#1 End-to-end hardware design, including deployment (hardware execution)
The design of the required computing units must be tailored to the hardware and the specific requirements of the application. Ideally, the behaviour of the actual sensors should also be considered. This highly specialised end-to-end hardware must be appropriately scaled using open-source toolchains on high-performance computers that includes the hardware influences (quantisation, type of computing units, activation approximations, computer architecture, etc.) of the specific hardware platform, map the entire data flow (from sensor to model) and be able to provide the translation to the platform. In this process, the optimised hardware design is created using a code generator and transferred to the hardware for execution. Ideally, there are no discrepancies between the optimisation on the computer and the execution on the final hardware.

#2 Dataset creation to enable an end-to-end AI system (optimisation, search space)
Real-world sensor data contains noise and may exhibit defects and environmental artefacts, which can significantly limit the performance of an AI system if these factors are not considered. In this context, (i) the selection of pre-processing methods (filtering, windowing, transformation, artefact detection and suppression, etc.), (ii) data encoding and adjustments, (iii) the capture of decision uncertainties (including incorrect labels in the dataset) and (iv) the detection and compensation of sensor drift data have a significant influence on the stability and accuracy of the AI system in the end application.

The presentation will highlight the challenges involved in designing a signal processing chain for neural tissue activity for use in a brain-computer interface (Fig. left) using digital and memristor hardware. By using memristors, an AI system with reconfigurable analogue computing can be built (Fig. right), which has the potential to outperform classical digital systems in terms of latency, resource usage and memory requirements.

[1] www.heise.de/en/news/ChatGPT-s-power-consumption-tentimesmore-than-Google-s-9852327.html
[2] de.ugreen.com/blogs/hausbatterie-backup/wie-viele-kwh-ein-haushalt-taglich-monatlich-und-jahrlich