Editorial

Reservoir computing with memristors

🔴 Check for updates

Memristive devices can be used to create energy-efficient hardware implementations of reservoir computing.

Artificial neural networks are at the heart of recent advances in artificial intelligence. The machine learning algorithms, which are inspired by biological neural networks and can find patterns in large amounts of data, have delivered a series of powerful applications in the past few years. Researchers at DeepMind in London have, for instance, developed a neural network (Alpha-Fold) that can predict protein structures[1] and, very recently, a neural network (AlphaTensor) that can discover algorithms for multiplying matrices[2]. Elsewhere, researchers at OpenAI in San Francisco have created a neural network language model (GPT-3) that is capable of generating sophisticated and fluent text[3].

The rapid increase in the performance of artificial neural networks can be linked to a rapid increase in their scale (GPT-3 has 175 billion parameters[4], for example). This though places considerable demands on the underlying electronic hardware — particularly their energy consumption. Conventional computing systems are based on a von Neumann architecture in which processing and memory units are physically separated and data have to be transferred between them. This architecture is inefficient when it comes to handling machine learning tasks. Thus, the search for new chip architectures and device technologies is increasingly critical.

Neuromorphic computing is one potential answer[5]. Emerging in the 1980s with the work of Carver Mead at the California Institute of Technology on analogue silicon technology, such brain-inspired systems now take a variety of different forms. A leading contender is neuromorphic computing based on memristive devices (or memristors), which
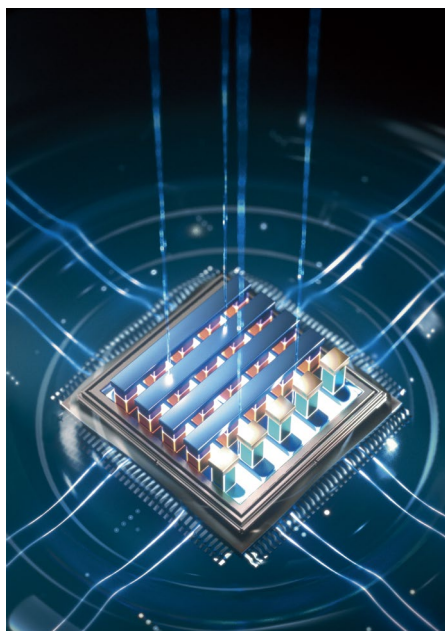
can provide both information processing and memory.

The artificial neural networks created with memristors are typically feedforward networks such as convolutional neural networks. These can struggle when it comes to handling temporal tasks. Recurrent neural networks, which have cyclic connections, can be used instead for such tasks. But training these networks can be challenging and implementing them with memristors problematic. Researchers have thus turned to an approach known as reservoir computing.

In reservoir computing, a dynamic 'reservoir' is used to map temporal inputs into a high-dimensional feature space. A readout layer is then used to analyse the extracted features. The reservoir is fixed and only the weights connecting to the final readout layer need to be trained (typically with a linear regression model). As a result, the technique can compute temporal data with a low training cost.

A variety of different components — from spintronic oscillators[6] to silicon modules[7] —

can be used to create hardware implementations of reservoir computing. But memristor-based systems have shown particular promise. For instance, dynamic memristors have previously been used to create reservoir computing hardware that is capable of efficient speech recognition and time-series forecasting[8]. However, these memristor-based systems are usually analogue–digital hybrids that have an analogue reservoir layer and a digital readout layer. An approach that incorporates both an analogue reservoir layer and an analogue readout layer would allow signals to be transmitted and processed through the system without any conversion.

In an Article in this issue of *Nature Electronics*, Jianshi Tang, Huaqiang Wu and colleagues report a fully analogue memristor-based reservoir computing system. (See also the News & Views article on the work from Xiaobing Yan at Hebei University.) The approach uses two types of memristor: dynamic memristors for the reservoir layer and non-volatile memristors for the readout layer. To illustrate the capabilities of the approach, the team — who are based at Tsinghua University, Soochow University and the University of Glasgow — show that it can complete temporal arrhythmia detection and spatiotemporal dynamic gesture recognition tasks with high accuracy. Furthermore, they highlight that the power consumption of the fully analogue approach is lower than digital hardware and hybrid digital–analogue systems.

References
1. Callaway, E. *Nature* **608**, 15–16 (2022).
2. Fawzi, A. et al. *Nature* **610**, 47–53 (2022).
3. Johnson, S. A.I. is mastering language. Should we trust what it says? *The New York Times* (15 April 2022); https://go.nature.com/3g1cbx5
4. Brown, T. B. et al. Preprint at https://doi.org/10.48550/arXiv.2005.14165 (2020).
5. *Nat. Electron.* **3**, 347 (2020).
6. Torrejon, J. et al. *Nature* **547**, 428–431 (2017).
7. Vandoorne, K. et al. *Nat. Commun.* **5**, 3541 (2014).
8. Moon, J. et al. *Nat. Electron.* **2**, 480–487 (2019).

CREDITS: JIANSHI TANG