

Self-rectifying memristors with high rectification ratio for attack-resilient autonomous driving systems

Received: 17 December 2024

Accepted: 8 June 2025

Published online: 01 July 2025

 Check for updates

Guobin Zhang^{1,2}, Xuemeng Fan^{1,2}, Jie Wang^{1,2}, Zijian Wang^{1,2}, Zhejia Zhang^{1,2}, Pengtao Li^{1,2}, Yitao Ma^{1,2}, Kejie Huang³, Bin Yu^{1,2}, Qing Wan⁴ ✉, Xiangshui Miao⁵ ✉ & Yishu Zhang^{1,2} ✉

With the rise of big data and the Internet of Things, smart devices, especially autonomous driving systems, have become prime targets for information leakage and cyberattacks. This study presents the design and fabrication of a self-rectifying memristor utilizing a TiN/HfO_x/Pt structure to enhance the security and reliability of autopilot systems. Following rapid thermal annealing treatment, the self-rectifying memristor demonstrates a recorded rectification ratio exceeding 10⁸ and a nonlinearity of over 10⁵, coupled with minimal device-to-device (3.32%) and cycle-to-cycle variations (1.55%). We further extend the application of self-rectifying memristors into crossbar arrays for the real-time classification of autonomous driving datasets, showcasing their capability to execute artificial neural networks at the hardware level. The proposed crossbar arrays exhibit robust attack resilience, achieving classification accuracy (84.25%) comparable to those of software models (84.34%), particularly under complex attack scenarios. This work not only highlights the potential of self-rectifying memristors in bolstering the security of autonomous driving systems but also offers innovative strategies for safeguarding future intelligent transportation systems.

The advent of big data and the Internet of Things (IoT) has transformed smart devices -- ranging from smartphones to autonomous driving systems (ADS) into high-risk targets for information leakage and cyberattacks^{1,2}. These devices, characterized by their wide distribution and diverse functionalities, handle massive amounts of edge data, making them vulnerable to malicious actors³⁻⁵. ADS, as a prominent example of IoT terminals, requires real-time processing of data from vehicle sensors and communication with other vehicles, infrastructure, and cloud services^{6,7}. However, the traditional von Neumann architecture, which separates storage and computation, necessitates frequent data transfers between storage and processing units. This separation not only increases processing time but also significantly

elevates energy consumption⁸, particularly in edge computing environments where instantaneous data processing is crucial. Inefficient data handling exposes end devices to prolonged vulnerabilities, heightening the risk of cyberattacks and jeopardizing the overall security and reliability of these systems⁹. To mitigate these challenges, in-memory computing architecture has been developed to fuse memory and computing functions within a single unit, enabling "in-place processing" of data¹⁰⁻¹².

Memristors have emerged as promising candidates for in-memory computing systems due to their advantageous properties, including high-speed reading/writing, multi-level storage capability, non-volatility, low power consumption, and low latency¹³⁻¹⁶. Unlike the traditional

¹College of Integrated Circuits, Zhejiang University, Hangzhou, PR China. ²ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou, PR China. ³College of Information Science & Electrical Engineering, Zhejiang University, Hangzhou, PR China. ⁴Yongjiang Laboratory, Ningbo, Zhejiang, China. ⁵School of Integrated Circuit, Huazhong University of Science and Technology, Wuhan, PR China. ✉e-mail: qing-wan@ytlab.ac.cn; miaoxs@hust.edu.cn; zhangyishu@zju.edu.cn

silicon devices in the von Neumann architecture, memristors enable the tight integration of data storage and computation through matrix-vector multiplication¹⁷, leveraging the physical principles of Ohm's law and Kirchhoff's law to facilitate massive parallel computing capabilities^{18–20}. However, the mass integration of memristors often encounters the issue of sneak path currents, where unintended current paths lead to misreading and crosstalk between array units, adversely affecting the accuracy of matrix-vector multiplication operations and, consequently, the training and classification performance of neuromorphic computing systems. To address this challenge, various configurations such as transistor-memristor (ITIM)⁵, diode-memristor (IDIM)²¹, and selector-memristor (ISIM)²² have been proposed. While these approaches can mitigate data distortion during read operations, they introduce complexities in circuit design and manufacturing, potentially increasing costs and reducing yield²³. In contrast, passive crossbar arrays composed of self-rectifying memristors (SRMs) offer a more elegant solution^{24–29}. By eliminating the need for additional switching elements, these arrays can prevent misinterpretation caused by crosstalk, thus enhancing power efficiency and scalability while providing a cost-effective solution for neuromorphic computing applications^{30–32}. However, a common issue with SRMs at this stage is that the RR is not large enough or RR, NL, and process ease cannot be accommodated simultaneously, which greatly limits the size of the array^{24,30}. Therefore, exploring SRMs with both superior RR, NL, and the required simple process flow can significantly advance the development of larger scale in-memory computing.

In this study, we design a SRM based on a simple TiN/HfO_x/Pt structure and extend its application into arrays for real-time classification of autonomous driving databases. We conduct comprehensive physical and chemical characterizations of the device's thin films. Following rapid thermal annealing (RTA), our SRMs achieve recorded high rectification ratios (RR) of 10⁸ and nonlinearity (NL) of 10⁵, along with minimal device-to-device (D2D) and cycle-to-cycle (C2C) variations. We elucidate the underlying conducting mechanisms and demonstrate the scalability of SRMs in crossbar arrays, achieving a proof-of-concept scalability exceeding 25.4 terabits. Furthermore, we validate the artificial synaptic properties of the SRMs, enabling the construction of hardware-based artificial neural networks for autonomous driving applications. Notably, we assess the performance of SRMs in training and classifying the Berkeley DeepDrive (BDD) dataset, revealing the attack-resilient accuracy of 32 × 32 crossbar arrays based on the proposed SRMs closely approaches that of the YOLOv9 model. This work highlights the potential of SRMs in enhancing the security of ADS and lays the groundwork for their broader application in edge IoT environments.

Results

Thin film characterization and analysis

The resistance state of a memristor is influenced by the history of electrical stimulation applied to the device, analogous to the function of synapses in the human brain. The resistance change in memristors can simulate synaptic weight changes, thereby mimicking the behavior of biological synapses³³. A schematic illustrating the synaptic function of the proposed single-layer SRM is presented in Fig. 1a. In biological systems, an input pulse signal from a presynaptic neuron triggers the release of neurotransmitters into the synaptic gap, which then migrate toward the postsynaptic membrane. Postsynaptic neurons integrate information by recognizing the intensity and timing differences between excitatory and inhibitory postsynaptic potentials from various synapses^{34,35}. This interconnected network of neurons forms a large-scale neural network. Similarly, a complete crossbar array of SRMs functions as an artificial synaptic system, where each SRM unit acts as a single synapse²³. When stimulated by voltage pulses along the word lines the SRMs yield a current response corresponding to the changes in their resistance state, which is then output along the bit

lines³⁴. The process effectively converts digital signals into analog signals.

The cross-sectional transmission electron microscopy (TEM) image of the proposed SRMs is shown in Fig. 1b, while the energy dispersive spectroscopy (EDS) images of platinum (Pt), hafnium (Hf), oxygen (O), nitrogen (N) and titanium (Ti) are presented in Fig. 1c. The thickness of HfO_x layer is maintained at 10 nm to ensure adequate insulation while balancing defect concentration. Using this structure, we fabricated the SRMs (Supplementary Fig. 1), with the size of 3 × 3 μm². Given that HfO_x significantly influences the electrical properties of the proposed memristors, it is crucial to analyze its chemical composition. The X-ray photoelectron spectroscopy (XPS) spectrum of the Hf element, shown in Fig. 1d, reveals two spin-orbit peaks corresponding to Hf 4f_{7/2} and Hf 4f_{5/2}, indicative of Hf⁴⁺ valence states in the pristine HfO_x. And the Hf 4f XPS spectrum of HfO_x with RTA is illustrated in Supplementary Fig. 2. The XPS spectrum of oxygen in pristine HfO_x (Fig. 1e) shows a coexistence of 38.65% lattice oxygen (O_L) with the peak at 531.03 eV and 61.35% oxygen vacancies (O_V) with the peak at 532.82 eV, confirming a substantial number of defects within the HfO_x films produced via magnetron sputtering. Previous studies indicate that a high concentration of O_V can hinder the self-rectification effect in memristors due to the formation of continuous conducting filaments^{36,37}. To mitigate this issue, we performed RTA on the HfO_x films at 150 °C for 600 s. The XPS spectrum of the oxygen following RTA (Fig. 1f) shows a significant reduction in O_V concentration to 18.11%, where the binding energy peak of O_L and O_V is 530.85 eV and 532.63 eV respectively, demonstrating effective repair of the vacancies. The variation in O_V concentration from the surface to a depth of 5 nm in the HfO_x films after the RTA process with the same parameters is illustrated in Supplementary Fig. 3. Additionally, to validate the appropriateness of the selected RTA parameters, Supplementary Fig. 4 presents O_V concentration data corresponding to various RTA temperatures and durations. Surface roughness is another critical factor influencing memristor performance³⁸. Figure 1g–j displays atomic force microscopy (AFM) images of HfO_x films randomly selected from four different fabrication batches, indicating stable and low root-mean-square (RMS) surface roughness. Smooth surfaces reduce contact resistance and optimize charge injection for more stable and repeatable resistive switching, while rough surfaces are prone to uneven electric fields, localized hot spots, and deformation, which disrupt switching uniformity³⁹. In SRMs, moderate roughness guides the uniform formation of continuous conductive paths consisting of oxygen vacancies within the HfO_x, enhancing uniformity, while excessive roughness leads to random dispersion of the paths, which worsens the cycle-to-cycle variations. In terms of storage properties, smooth surfaces provide a stable environment for the conductive pathways within the HfO_x, prolonging retention time, whereas bad roughness accelerates film degradation, shortening retention and further weakening reliability. Endurance is also regulated by roughness, with optimal parameters enabling stable filament cycling, while excessive roughness triggers conductive filament destabilization, dramatically reducing the number of reliable switching³⁹. The surface profile data is presented in Supplementary Fig. 5, further confirming that the thin-film preparation process employed in this work ensures high-quality film formation of HfO_x.

Electrical characterization and scalability simulation

After thoroughly characterizing the HfO_x films from both physical and chemical perspectives, we proceeded to evaluate the electrical properties of the SRMs. The direct current (DC) current-voltage (I-V) curves of the SRMs and their wiring configuration during testing are illustrated in Fig. 2a. In this setup, the Pt bottom electrode (BE) is biased while the TiN top electrode (TE) is grounded. The choice of operating voltage scheme critically influences the performance of the SRMs; we selected the 1/6 V scheme (Supplementary Fig. 7) to optimize RR and

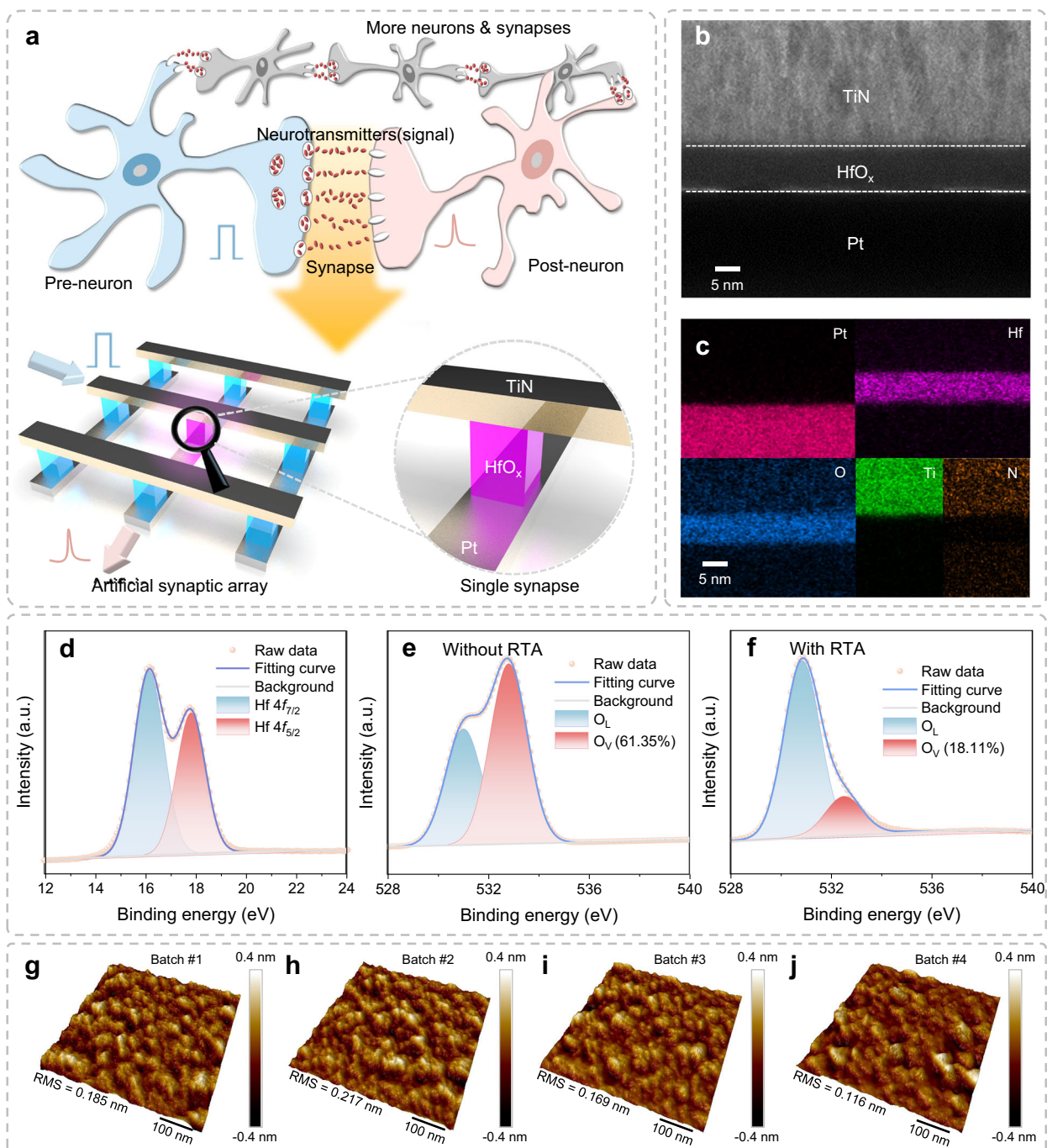


Fig. 1 | Physical and chemical characterization of thin films. **a** Scheme for artificial synapses based on the SRMs with the structure of TiN/HfO_x/Pt. **b** The TEM image of the cross-section of SRMs with the structure of TiN/HfO_x/Pt. **c** EDS

mapping images for Pt, Hf, O, Ti. **d–f** The XPS spectra of Hf, O without RTA, and O with RTA under 150 °C for 600 s. **g–j** AFM maps of HfO_x surfaces for four manufacturing batches.

NL²⁷. RR is defined as the ratio of the low-resistance state (LRS) current at the positive read voltage to that at 2/3 negative read voltage ($I_V/I_{-2/3V}$), while NL is defined as the ratio of the LRS current at the positive read voltage to that at the read voltage of a partially selected cell ($I_V/I_{1/6V}$) (Supplementary Fig. 7)²⁶. The proposed single-layer SRMs achieve a recorded RR of 10^8 and NL of 10^5 at the 1/6 V scheme, with a negative leakage current suppressed below 0.1 pA. Additional trends of RR and NL at varying operating voltages can be found in Supplementary Fig. 8. To assess the consistency among devices within the fabricated 32×32 array, we present the DC electrical characteristics of 16 randomly selected SRM units from the array in Supplementary

Fig. 10, confirming negligible variations. The significance of the RTA process is highlighted in Supplementary Fig. 13 and 14, which shows the I-V curves of TiN/HfO_x/Pt memristors without and with the RTA process at different parameters of RTA, demonstrating these parameters (150 °C, 600 s) are optimal for achieving self-rectification. Critically, the absence of self-rectification in non-RTA devices aligns with their filament-dominated conduction, as evidenced by the low-resistance Ohmic I-V profile. After the SET process, the conductive filament-based memristor is in the LRS. However, the robust conductive filaments lead to a negative RESET process that needs to undergo a scanning current with a very high peak value, which in turn

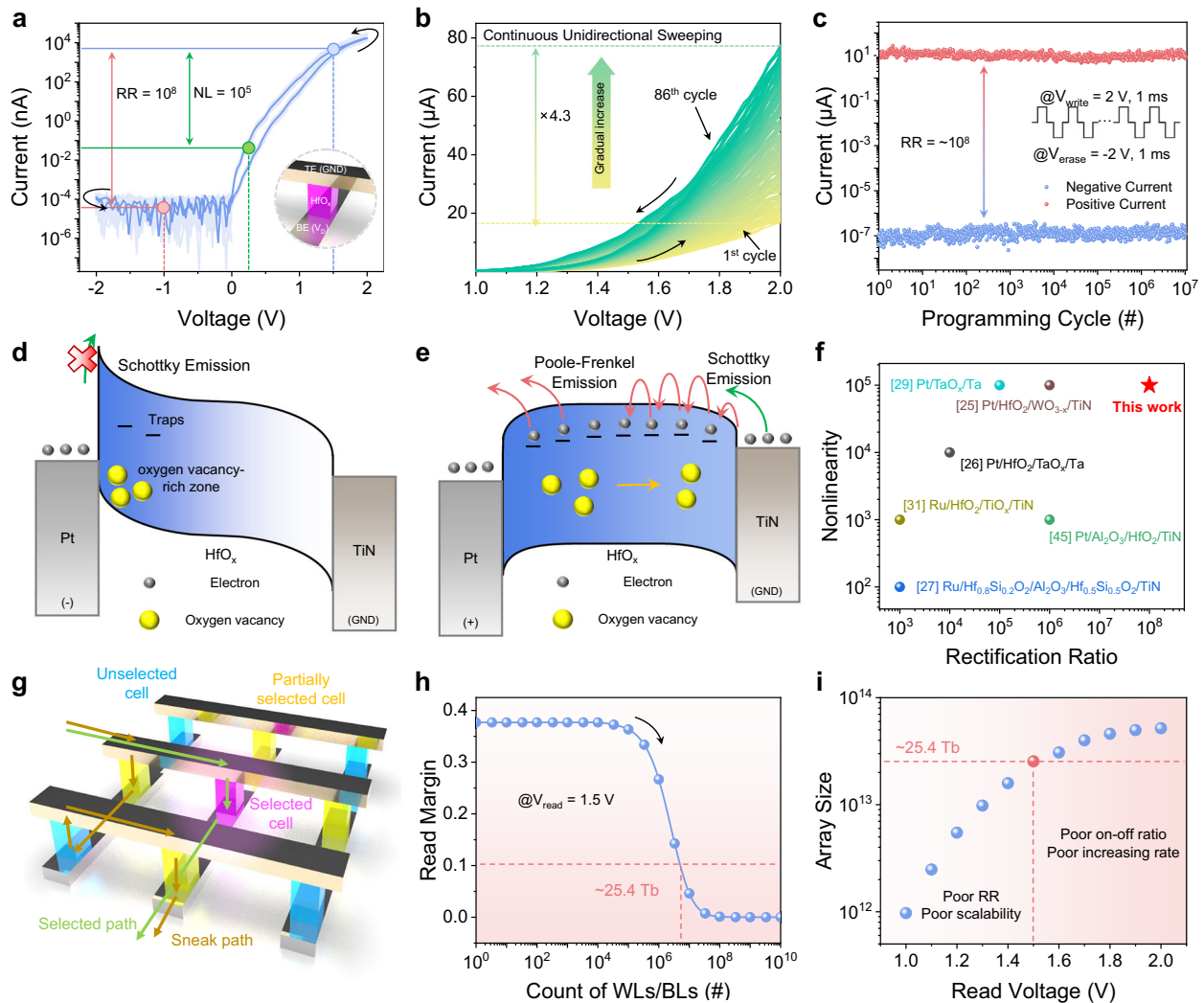


Fig. 2 | Electrical properties and scalability of the proposed SRMs. **a** I-V curves of TiN/HfO_x/Pt SRMs based on 1/6 V scheme with the RR of -10^8 and the NL of 10^5 . **b** Dynamic I-V curves obtained from the proposed SRMs with continuous unidirectional sweeping. **c** RR endurance properties of the proposed SRMs over 10^7 programming cycles. The conducting mechanisms of the proposed SRMs when

d negative bias and **e** positive bias is added respectively. **f** A comparison chart of this work with advanced SRMs compatible with CMOS technology on two important metrics: RR and NL. **g** Schematic diagram of sneak path currents in the crossbar array without self-rectifying. **h** The scalability simulation of the proposed SRMs with over 25.4 Tb scale. **i** The trend chart of array size with DC operating voltage.

makes the self-rectification effect non-existent and thus poorly scalable⁴⁰. In contrast, RTA-treated devices exhibit nonlinear, rectifying behavior, which correlates with suppressed filament formation. This transition is further supported by the quantified O_v concentration gradient derived from XPS depth profiling. When the RTA temperature or time is insufficient, the O_v distribution inside the HfO_x is uniform, and the conducting filaments can be easily formed and thus realize the abrupt SET⁴⁰. Figure 2b illustrates the response of the SRMs following 86 consecutive positive voltage sweeps from 0 to 2 V, revealing a gradual increase in response current corresponding to a continuous increase in conductance. This consistent change in conductance effectively mimics the dynamics of synaptic weights, affirming the device's capability to replicate potentiation behaviors observed in biological synapses. The RR endurance performance of the SRMs is characterized in Fig. 2c, where repeated programming with alternating positive pulses (2 V/1 ms) and negative pulses (−2 V/1 ms) results in minor fluctuations in response current, indicating excellent stability in conductance and RR over more than 10^7 programming cycles. In addition, we tested the endurance characteristics of the on/off ratio of this SRM (Supplementary Fig. 15), demonstrating the reliability of the

proposed single-layer SRMs for long-term repeated programming. Moreover, we tested the device for data retention and obtained the conductance decay curves shown in Supplementary Fig. 16.

We further elucidate the conducting mechanism responsible for the ultra-high RR of the SRMs. The selection of TiN and Pt as TE and BE, respectively, is based on their work functions of −4.5 eV and 5.6 eV^{41,42}, leading to a significant asymmetry in the energy band structure (Supplementary Fig. 17). After the RTA process, a substantial number of O_v in the HfO_x layer near the TiN/HfO_x interface are repaired, while a similar concentration remains at the Pt/HfO_x interface. Under negative bias (Fig. 2d), O_v migrates in the direction of the electric field, accumulating near the Pt/HfO_x interface and resulting in an unbalanced O_v distribution. As negative voltage increases, electrons from the BE can't undergo Schottky emission on account of the large barrier at Pt/HfO_x interface, resulting in negative rectification⁴³. Conversely, when a positive bias is applied (Fig. 2e), O_v migrates toward the TiN/HfO_x interface, leading to a more uniform distribution and the formation of a continuous electron-conducting pathway. As the positive bias increases, electrons undergo both Schottky and Poole-Frenkel emissions, facilitating significant current response. Under positive bias, the

HRS is characterized by sparse, non-uniform oxygen vacancy clusters forming discontinuous percolation paths before SET, while the LRS emerges through the formation of continuous, dense, and homogeneous conductive paths driven by field-enhanced vacancy migration after SET (Supplementary Fig. 18). The simulation results shown in Supplementary Fig. 19 confirm the migration as well as the distribution of O_V in HfO_x with the electric field as previously described, and the detailed scheme is described in the “Methods” Section. To further investigate the impact of TE materials, we also utilized indium tin oxide and tungsten (W), which have similar work functions to TiN^{44,45}. The resulting I-V curves (Supplementary Fig. 20) indicate the DC properties of the proposed SRMs are not directly influenced by the chemical composition of the TEs. The electron affinity of HfO_x and the metal's work function are the key determinants of the metal/oxide Schottky barrier. Supplementary Fig. 21 demonstrates the electron affinity of the upper surface of the HfO_x without RTA as obtained by Low Energy Inverse Photoelectron Spectroscopy (LEIPS) characterization⁴⁶. The value is 3.542 eV in the absence of an applied electric field, which is closer to the figure of merit of 4.5 eV for TiN. And when positive bias is applied, as mentioned before, the Schottky barrier of TiN/ HfO_x decreases significantly due to the increased electron affinity of that side of the HfO_x , resulting in significant Schottky emission at low voltage with guaranteeing the large enough barrier of Pt/ HfO_x . Furthermore, in order to verify the important influence of Schottky emission and P-F emission on the conductive mechanism of the proposed SRM, we also tested the DC I-V curves at different high temperatures and fitted them linearly (Supplementary Fig. 22). Further, based on the data at high temperatures, the variation of the Schottky barrier at the TiN/ HfO_x interface with temperature is illustrated in Supplementary Fig. 22, which is 0.16 eV at low temperatures and 0.56 eV at high temperatures⁴⁷. Notably, the complementary metal-oxide semiconductor (CMOS)-compatible SRMs demonstrated superior performance (Fig. 2f), achieving optimal RR and NL values with scalability to date^{24–32,48}. The single-layer structure of the proposed SRMs is simpler and more straightforward to fabricate than conventional two-layer or multilayer designs, and HfO_x is compatible with CMOS technology, enabling potential large-scale integration on silicon-based chips^{49,50}.

Based on the single-layer SRMs, we constructed 32×32 crossbar arrays to study their scalability and explore their feasibility as a hardware platform for neuromorphic computation. Figure 2g presents a schematic of the sneak paths in localized 3×3 sections of the 32×32 crossbar arrays. Ideally, only the selected cell should conduct current, while unselected cells remain inactive. The equivalent circuit diagram of the sneak path in the passive crossbar array is shown in Supplementary Fig. 23. Inter-cell misreading occurs only when current flows through the selected cell. However, both SRMs and memristors without self-rectification exhibit pronounced inter-device crosstalk when scaled to larger array sizes⁵¹. We employed a pull-up voltage strategy (detailed in Supplementary Note 1) to calculate the read margin and simulate the maximum scalability achievable with the proposed SRMs⁵¹. Increasing the read margin allows for a larger operating voltage window, improving the distinction between the LRS and high resistance state (HRS) of the targeted cells during the read operations. However, the read margin typically decreases as the crossbar array size increases. It is generally accepted that when the read margins drop below 10% in the worst case when all unselected cells are incorrectly activated or read, the maximum realizable array size is reached³⁰. As shown in Fig. 2h, the single-layer SRMs achieve scalability exceeding 25.4 Tb at a read voltage of 1.5 V, representing state-of-the-art performance (Supplementary Table 1). We also examined the factors affecting the scalability of the proposed SRMs at different read voltages. Results in Fig. 2i indicate that read voltage below 1.5 V results in poor RR and NL (as shown in Supplementary Fig. 8), leading to limited scalability. Conversely, voltages larger than 1.5 V yield a poor on/off ratio in the I-V curve, similarly constraining scalability. Thus, the

combination of scalability, voltage magnitude, and reliability establishes 1.5 V as the optimal read voltage, ensuring scalability without compromise.

Characterization of artificial synaptic properties

As shown in Fig. 2b, the proposed SRMs exhibit a weight potentiation analogous to that of biological synapses. In this section, we demonstrate weight updating through electrical pulse testing. We first applied a positive writing pulse and a negative RESET pulse to a single SRM within the 32×32 array, ensuring both pulses had the same width, amplitude, and adjustable intervals. Figure 3a depicts the applied electrical pulse (blue) alongside the corresponding current response (red). Under the positive writing pulse, the device generates a peak current of about 1.3 μ A. In contrast, the negative RESET pulse elicits minimal current output due to the previously described rectification effect. For subsequent synaptic weight update tests, we sampled the response current at the onset of the falling edge of the positive pulse. Paired pulse facilitation (PPF) is a fundamental characteristic of biological synapses, reflecting the enhanced response to successive stimuli⁵². PPF serves as an indicator of synaptic plasticity, which is crucial for artificial synapses in simulating learning and memory processes. The variation of the PPF index with pulse interval time is shown in Fig. 3b, with the PPF index defined by Eq. 1:

$$PPF \text{ index} = \frac{I_2 - I_1}{I_1} \times 100\% \quad (1)$$

where I_1 is the initial peak current response and I_2 is the subsequent peak current response. With a fixed pulse width of 0.2 ms, the PPF index exhibits an exponential decline with increasing interval time, stabilizing at 0 after 200 ms. The equation for PPF versus interval time t is presented by Eq. 2, where t_1 and t_2 are the decaying coefficient (approximately equal to 7.44 and 67.88, respectively) as well as A_1 and A_2 are both constant.

$$PPF \text{ index} = A_1 \exp\left(-\frac{t}{t_1}\right) + A_2 \exp\left(-\frac{t}{t_2}\right) \quad (2)$$

Due to the rectification effects, the statistics of paired-pulse suppression are applicable in this context. Synaptic plasticity in biological systems manifests in two primary forms: long-term potentiation (LTP) and long-term depression (LTD). To further investigate these phenomena, we fixed the interval time at 0.2 ms, and explored the dynamic range and linearity of weight updates across various pulse widths (Fig. 3c). Five pulse widths were compared: 20 μ s, 0.1 ms, 0.2 ms, 1.0 ms, and 2.0 ms. The results indicate that longer pulse widths (from 0.1 ms to 2.0 ms) exhibit decreased noise relative to 20 μ s pulse width, as inadequate linearity in artificial synaptic updates can severely impair neural network training accuracy and lead to misclassification. At the same time, according to LTD, longer pulse widths lead to a severe weakening of the self-rectifying effect thus worsening the misreading phenomenon in the crossbar array. Conversely, the 20 μ s pulse width keeps the self-rectifying effect with minimal noise and great updating linearity. Notably, unlike conventional O_V conducting filament-based LTD³⁴, the LTD observed here is constrained by the rectification effect²³.

Figure 3d presents the details of LTP and LTD for the artificial synapses based on the proposed SRMs. The subtle C2C variations in weight updating of memristors can significantly interfere with data processing⁵³, making minimal D2D and C2C variations essential for enhancing the accuracy of neuromorphic computations. Figure 3e displays the LTP and LTD characteristics across the 100th, 200th, 300th, 400th, and 500th cycles, demonstrating exceptionally small C2C variations. The main purpose of the LTD in Fig. 3d, e is to ensure

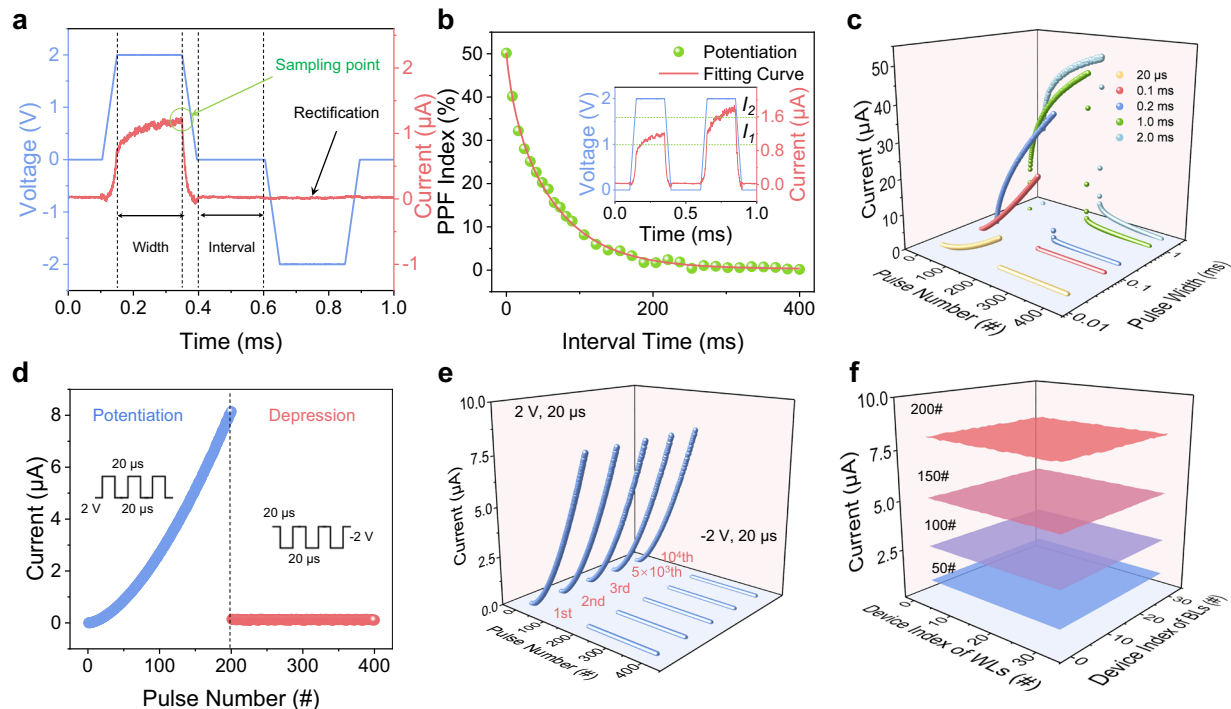


Fig. 3 | Artificial synaptic properties of the proposed SRMs. **a** One write pulse as well as one RESET pulse (blue) and the resulting response current (red). **b** The variation of PPF with the paired voltage pulse interval. Inset: Schematic diagram of the measure of PPF. **c** LTP and LTD for weight updating of proposed SRMs at

different voltage pulse intervals. **d** LTP (blue) and LTD (red) for weight updating of proposed SRMs at the pulse interval of 20 μ s. **e** Cycle-to-cycle variations of weight updating at the pulse interval of 20 μ s. **f** Device-to-device variations of weight updating in one 32×32 crossbar array.

that the rectification effect exists and thus the sneak path currents are still effectively suppressed when scaling to larger arrays²³. In order to visualize the symmetric LTP/LTD, we tested 10^4 consecutive LTP/LTD cycles illustrating the excellent endurance of this artificial synapse (Supplementary Fig. 24). We further characterized the randomness of LTP across all SRM cells in a 32×32 crossbar array (Fig. 3f). Supplementary Figs. 25 and 26 provide comprehensive statistics on the distribution of 200 LTP simulation states for 1024 SRMs, where there is no overlap between the different states in the crossbar array. The results indicate that the D2D variations become increasingly significant with a higher number of applied pulses. Statistical analysis of D2D and C2C variations for all SRMs in the array is presented in Supplementary Fig. 27. We conducted statistical analyses across 1024 devices (D2D variations) and 10^4 consecutive LTP cycles (C2C variations) to validate the reproducibility. The low D2D variations of 3.32% and C2C variations of 1.55% confirm that the interfacial switching mechanism, combined with process control, mitigates the variability commonly observed in filamentary memristors^{6,38}, ensuring the ultra-high weight-updating precision of the proposed SRMs for subsequent neuromorphic applications.

ADS based on SRMs

Based on the super-optimal self-rectifying function along with the excellent artificial synaptic properties of SRMs, we further verify the feasibility of applying SRMs to intelligent recognition in ADS. The driving of the SRMs-based crossbar arrays was realized by a printed circuit board (PCB) as shown in Fig. 4a, where Fig. 4b, c are the local zoomed-in chip carrier and SEM image of the 32×32 array on the chip carrier, respectively. The peripheral circuitry driving the SRM-based passive crossbar array is shown in Supplementary Fig. 28. We employed the YOLOv5 (Supplementary Fig. 29) and YOLOv9 (Supplementary Fig. 30) models, which are prominent single-stage target detection algorithms. YOLOv9, the latest version, offers significant

improvements in efficiency and accuracy compared to YOLOv5 through programmable gradient information and an efficient layer-aggregation network architecture (Fig. 4d). The network structure consists of three components: The Backbone network for feature extraction, the Neck network for enhancing fused features, and the Head network for classification. To validate the combined hardware and software models for real-time traffic conditions, we trained both models using a comprehensive dataset, including the COCO 2017 Dataset, Mapillary Traffic Sign Dataset, LISA Traffic Sign Dataset, Lyft L5 Dataset, and 80% of the Berkeley DeepDrive Dataset (BDD100K)⁵⁴. We allocated 10% of the BDD100K dataset for validation and the remaining 10% for testing, unifying labels into five categories: pedestrians, vehicles, signposts, obstacles, and stoplights (Supplementary Fig. 31). The connection weights in the neuromorphic architecture (Fig. 4d) were updated in real time using the SRM-based crossbar array.

After training, we obtained classification accuracies for the BDD100K dataset, as shown in Fig. 4e–j. The first two columns of the histograms represent the software-based YOLOv5 and YOLOv9 models, while the last two columns correspond to the memristors-based YOLOv5 (M-YOLOv5) and YOLOv9 (M-YOLOv9). All models achieved real-time classification accuracies exceeding 90% for five labels, closely aligning with the requirements for ADS. Notably, the SRM-based models outperformed their software counterparts in stoplight classification due to minimal variations and 200 instantaneous analog states, enhancing weight-updating accuracy. The crosstalk-free crossbar array also aids in feature extraction and noise reduction⁵³. The classification accuracies for green and red lights are detailed in Supplementary Fig. 32. As shown in Fig. 4j, the accuracies for M-YOLOv5 and M-YOLOv9 are nearly identical to those of the software model, differing by only 0.09% and 0.06%, respectively. The SRM-based algorithm offers distinct advantages over conventional YOLOv5 and YOLOv9 architectures in ADS, primarily through its intrinsic analog computing paradigm and hardware-algorithm co-design. Unlike static

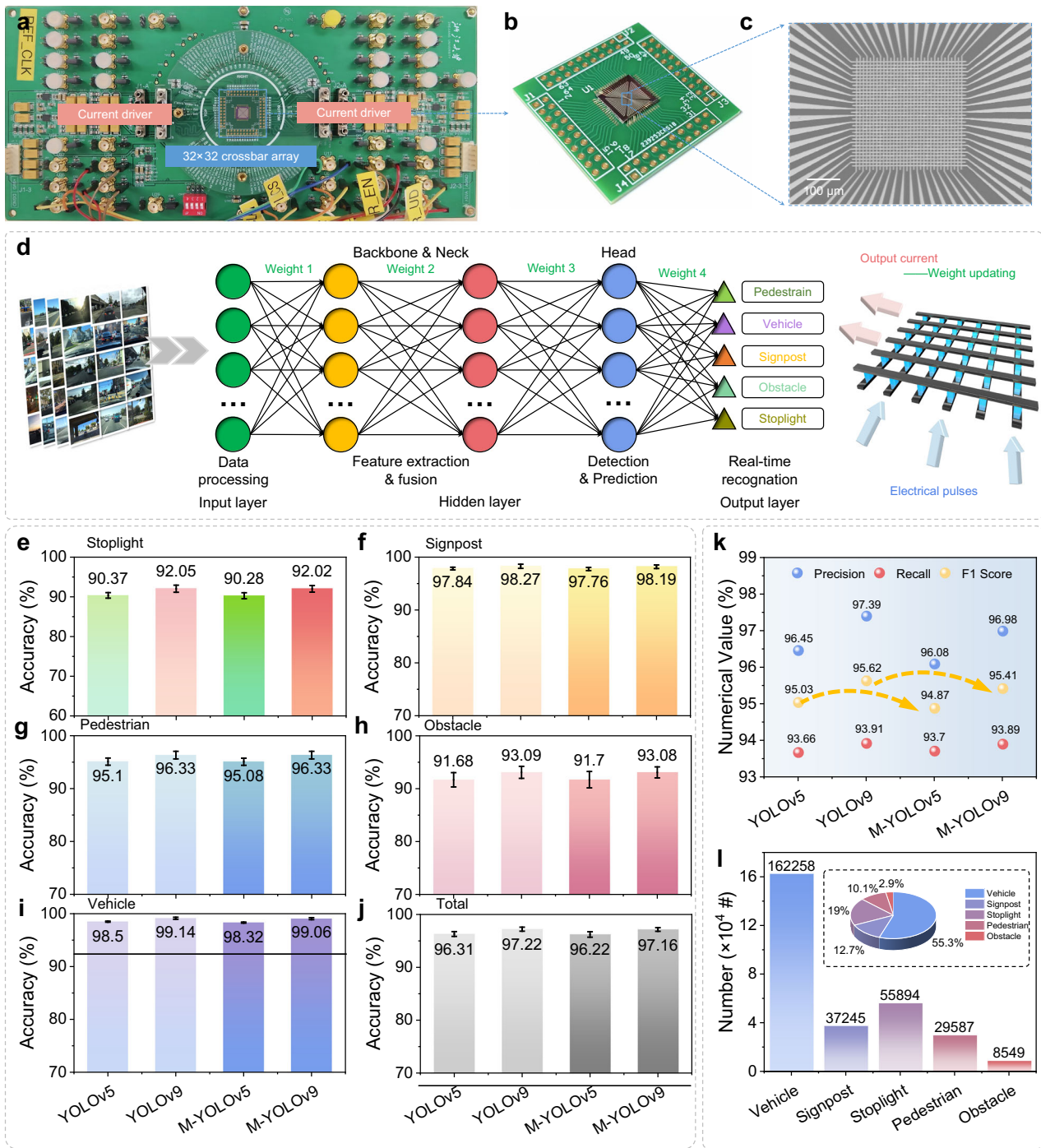


Fig. 4 | Intelligent recognition system for ADS based on SRMs. a PCB board integrating SRMs-based 32×32 crossbar array and driver circuits. **b** Chip carrier with 32×32 crossbar array encapsulated on the PCB board. **c** SEM image of the 32×32 crossbar array. **d** Neuromorphic model architecture of YOLOv5 and YOLOv9 for the BDD dataset. The classification accuracy of **e** stoplight, **f** signpost, **g** pedestrian, **h** obstacle, **i** vehicle, and **j** total aspects based on the four models:

YOLOv5, YOLOv9, M-YOLOv5, and M-YOLOv9, where the sample size is 100. The error bars represent the maximum and minimum intervals of the obtained batch results. **k** The precision, recall, and F1 score obtained by the four models YOLOv5, YOLOv9, M-YOLOv5, and M-YOLOv9 on the BDD dataset. **l** The total count of five labels—stoplight, signpost, pedestrian, obstacle, and vehicle in the BDD dataset.

YOLO models, which rely on fixed-weight inference post-training⁵⁵, the SRM's analog tuning capability enables dynamic feature refinement during real-time operation. This is achieved via sub-precision, context-aware weight adjustments that adaptively recalibrate traffic light detectors (e.g., color and shape recognition) under challenging conditions such as fluctuating illumination or partial occlusion—a scenario where traditional YOLO architectures may exhibit suboptimal

robustness due to their rigid parameterization⁵⁶. Furthermore, the analog memristive network circumvents quantization artifacts inherent in digital implementations, which often degrade nuanced feature representations critical for discriminating fine-grained traffic light states (e.g., distinguishing red and yellow signals under glare or fog)^{57,58}. By preserving analog feature fidelity, the SRM-based system mitigates misclassification risks associated with amplitude-truncated

digital activations. Additionally, the spatially constrained plasticity of memristive arrays introduces localized learning mechanisms, where weight updates are influenced predominantly by active input regions rather than global network perturbations. This localized adaptation enhances feature-specific learning efficiency, particularly in class-imbalanced traffic scenarios where rare but safety-critical events (e.g., malfunctioning traffic lights) require prioritized representation without destabilizing the broader model⁵⁹.

We further evaluated model performance using four metrics: precision, recall, and F1 score, which are critical for assessing efficacy. Precision is of central value in reducing false positives and improving decision quality, recall ensures that critical information is not missed, and F1 score quantifies the optimal trade-off point between precision and recall. Detailed metrics are provided in Supplementary Note 2. As depicted in Fig. 4k, the SRMs-based YOLOv5 and YOLOv9 models exhibit comparable metrics to their software-based counterparts, confirming the compatibility of the proposed SRMs with YOLOv9. Additionally, we quantified the distribution of labels in the BDD100K dataset (Fig. 4l). Vehicles account for over 50% of the dataset, while the remaining four labels collectively represent 45%. This distribution indicates that the model's accuracy for vehicles significantly influences overall classification accuracy. Besides, we added comprehensive experimental results (Supplementary Fig. 33) to explicitly illustrate the effect of pulse size (1 V, 1.5 V, 2 V, 2.5 V, 3 V) and the corresponding LTP/LTD behaviors on ADS recognition accuracy. The results show that the recognition accuracies for all four non-optimal pulse intensities (1 V, 1.5 V, 2.5 V, 3 V) are significantly lower compared to the 2 V benchmark due to negative factors such as high asymmetry and poor linearity in important weight updating.

Anti-attack characteristics of ADS

The feasibility of the proposed SRMs combined with YOLOv5 and YOLOv9 models has been verified. In the IoT era, addressing hardware design cost and power consumption while ensuring the stability of ADS is critical. The SRM arrays reduce hardware size and perform neuromorphic computation with low power consumption, comparable to GPUs²⁴, enhancing the resilience of ADS against attacks from algorithmic or physical sources. Figure 5a illustrates the flow of using SRM arrays to enhance ADS. The vehicle's camera transmits real-time road video to the SRM array, which minimizes data transmission delays and accelerates decision-making. The refined process is detailed in Fig. 5b, where the details are demonstrated in the Methods module.

To simulate attacks on the vehicle's camera, we applied six types of strains to BDD100K images, including blurry patches and RGB patches (Fig. 5c). Other examples of different attack patterns are presented in Supplementary Figs. 34–40. The patch attack concept involves covering random image positions with ellipse-shaped patches and the number of patches, the transparency maximum (α_{\max}), are adjustable. Figure 5d shows classification accuracy for the SRM-based model, GPU-based model, and non-self-rectifying memristors (NSRM) model under these attacks. The YOLOv9 model based on NSRM exhibits significantly reduced accuracy due to poor weight-updating precision. NSRMs (the first figure of Supplementary Fig. 13) exhibit poor robustness under attack conditions due to their reliance on filament-dominated conduction⁴⁰ and the absence of leakage current suppression, suffering from uncontrolled sneak path currents during adversarial perturbations, which amplify noise propagation across the array. This results in feature ambiguity and quantization error accumulation. Attacking patches exploit the lack of spatial selectivity for NSRM-based ADS, causing unintended activation of adjacent synapses and blurring decision boundaries. And digital-to-analog conversion bottlenecks along with accumulated errors in NSRM arrays exacerbate adversarial sensitivity. In contrast, the SRM-based model maintains accuracy comparable to the GPU-based model across all attack types, with RGB patches yielding optimal results. The effects of varying

transparency and patch numbers on recognition accuracy are detailed in Supplementary Figs. 41 and 42.

To further evaluate model performance, we analyzed precision and recall using the precision-recall (PR) curve, which illustrates the trade-off between these metrics at different decision thresholds. A well-performing model should have a PR curve that is distant from the diagonal and close to the upper right corner. As shown in Fig. 5e, the SRM-based model's PR curve closely matches that of the GPU-based model under RGB patch attacks, while the NSRM implementation shows poor performance. We also simulated the impact of D2D and C2C variations on the recognition accuracy. Figures 5f, g indicate that the SRM arrays can tolerate D2D variations up to 10% without significant accuracy lost, while tolerance for C2C variations is limited to about 2%. The ultra-low variations in the proposed SRM arrays are effectively managed through self-correction during weight updates, ensuring robust recognition accuracy. Finally, Fig. 5h presents a schematic diagram of the four decisions made in response to an RGB-patch attack, while the confusion matrix in Fig. 5i illustrates the strong real-time identification and decision-making capabilities of the SRMs-based YOLOv9 model. Moreover, since the proposed SRMs are quasi-volatile, and also the demand for weight retention capability of each cell in the array is high in the hardware-based YOLO model inference process⁶⁰, we comprehensively calculate the approximate energy consumption required for refreshing and reloading the weights in Supplementary Note 3.

Discussion

In this work, we successfully designed and fabricated a SRM with an ultra-high RR and a simple structure, demonstrating its potential as an attack-resilient technology for autonomous driving. The TiN/HfO_x/Pt-based SRM achieves over 10⁸ RR and 10⁵ nonlinearity (NL) after RTA. Experimental results show that the SRM-based ADS achieves classification accuracy comparable to software models, thanks to minimal variations and 200 artificial synaptic states that optimize weight-updating accuracy. Additionally, the SRM maintains an attack-resilient accuracy similar to GPU models across various attack scenarios, underscoring its resilience against algorithmic and physical threats. This SRM technology is poised for widespread application in autonomous driving, promising to advance intelligent transportation systems through its high speed, low power consumption, and enhanced security.

Methods

Preparation of thin films

In this work, all films were prepared using a magnetron sputtering system (DISCOVERY-635, Denton Vacuum) with meticulously controlled parameters to ensure reproducibility. Initially, a 50 nm Pt bottom electrode was sputtered onto SiO₂/Si substrates pre-patterned via photolithography (AZ5214 photoresist, 1.2 μm thickness). The Pt deposition was performed at a DC power of 300 W with an Ar gas flow of 37 sccm, a base pressure of 5.0 × 10⁻⁶ Torr, and a working pressure of 3.0 mTorr. The substrate temperature was maintained at 25 °C (unheated) with active water cooling to minimize thermal stress, and the deposition rate was calibrated to 0.71 nm/s (70 s duration). Following Pt deposition, the photoresist was removed using acetone and isopropanol ultrasonic baths, followed by N₂ drying. Next, a 10 nm HfO_x switching layer was reactively sputtered from a HfO₂ target (99.995% purity) in the pure Ar atmosphere with the gas flow of 37 sccm. The substrate temperature was actively controlled at 50 °C ± 5 °C to optimize film uniformity and O_v density, and the RF power was set to 50 W with a deposition rate of 0.013 nm/s (770 s duration). Post-deposition, the HfO_x films underwent RTA in an RTP-VT100M system (ULVAC) under O₂ ambient (99.999% purity). The annealing profile included a ramp rate of 25 °C/s, a peak temperature of 150 °C, a dwell time of 600 s, and natural cooling to <50 °C, with the

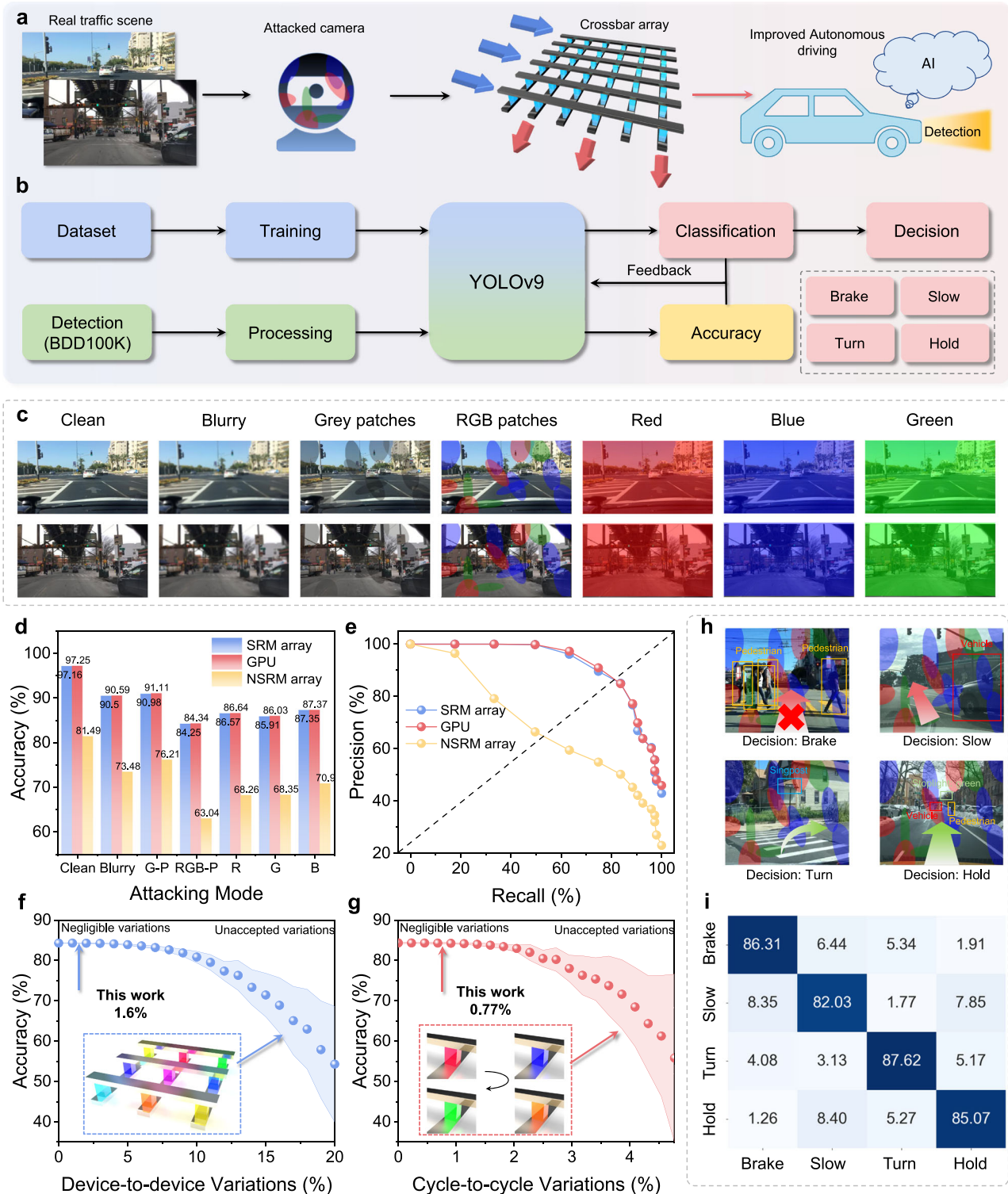


Fig. 5 | Anti-attack tests for ADS based on SRMs. **a** Schematic diagram of the hardware-software cooperative anti-attack ADS based on the proposed single-layer SRMs. **b** Flowchart of real-time classification and decision making for YOLOv9 models combined with SRM arrays. **c** Sample images of the original traffic dataset with the six attack patterns. **d** Classification accuracy of YOLOv9 model combined with SRM array, GPU-YOLOv9 model, and YOLOv9 model combined with NSRM array on the BDD100K dataset after six-pattern attacks, where G-P denotes Gray Patches, RGB-P denotes RGB Patches, and R, G, B refers to red, green, and blue,

respectively. **e** PR curves of YOLOv9 model combined with SRM array, GPU-YOLOv9 model, and YOLOv9 model combined with NSRM array on the BDD100K dataset after RGB-patch attacks. Variation of classification accuracy of SRMs-based YOLOv9 model with different **f** device-to-device variations (different colors correspond to different device states), and different **g** cycle-to-cycle variations (different colors correspond to state changes of the same device). **h** Four paradigm scenarios of decision from ADS. **i** Confusing matrix of four aspects of decision from ADS based on SRMs.

ambient pressure maintained at 10 mTorr to suppress oxygen exchange and stabilize the O_v gradient. Subsequently, a 50 nm TiN top electrode was sputtered using a TiN target (99.995% purity) with an Ar gas flow of 37 sccm. The substrate temperature was set to 25 °C (unheated), and the DC power was 200 W with a deposition rate of 0.21 nm/s (240 s duration). Finally, all remaining photoresist was removed via sequential acetone (10 min) and isopropanol (5 min) ultrasonic baths, followed by N_2 drying.

Characterization of thin films

High-resolution TEM images of cross-sections were obtained using a Thermo Scientific Helios G4 HX, with samples prepared via the in-situ focused ion beam (FIB) lift-out technique. To protect the sample surface from ion-beam damage during extraction and thinning, electron-beam-deposited platinum (e-PT) and ion-beam deposited platinum (i-Pt) were applied. EDS was conducted using a Super X FEI System to analyze all elemental compositions. Additionally, XPS for the elements of Hf and oxygen O was performed with a Thermo Scientific ESCALAB 250Xi model, with an Al Ka (1486.6 eV) XPS source. AFM characterization was carried out using a Dimension ICON instrument. The AFM images shown in Fig. 1g-j were all obtained from four batches of prepared HfO_x films subjected to RTA at 150 °C for 600 s, after the sputtering processes utilizing a power of 50 W for 400 s at a chamber pressure of 3.72 mTorr. UPS and LEIPS were tested with an E03-001 Scanning Focusing Multi-Platform XPS system, model PHI 5000 VersaProbe, with an XPS energy resolution of 0.1 eV and a filtering energy of 4.38 eV. The parameters of UPS and calibration parameters for the LEIPS pre-test included an X-ray source with a HeI 21.22 eV at 80 W and a beam spot diameter of 5 μ m.

Defect migration simulation

The finite element analysis simulations are performed with the COMSOL Multiphysics software. The FEA model in shown below was defined by the law of mass conservation and electroneutrality assumption of related ions without considering the possible side reactions. A 2D model with dimensions of $7 \times 5 \mu$ m was created in the simulation model, and the effects of ion concentration diffusion and ion migration on ion concentration were considered. The simulation model considers three different models: (1) No potential influence; (2) Electric field distribution from bottom to top; (3) Electric field distribution from top to bottom. By setting up these three models, the influence of the electric field direction on ion migration and diffusion distribution is studied. The diffusion coefficient of each domain is determined based on the relationship between the ionic conductivity and the diffusion coefficient in the Nernst-Einstein equation. Therefore, under the combined action of electric field migration and concentration diffusion, the anions and cations move in different directions, eventually reaching a steady state.

Using the current distribution and concentration diffusion equation to track the current and concentration distribution, the flux of each ion in the electrolyte can be calculated by Nernst-Planck equation:

$$N_i = -D_i \left(\nabla c_i + \frac{F}{RT} z_i c_i \nabla \phi_l \right), i = 1, 2, \dots, n \quad (3)$$

Where D_i , N_i , c_i , z_i , F , T , R , and ϕ_l denote the diffusion coefficient, the flux, the concentration, the charge number, the Faradaic constant, temperature, gas constant, and electrolyte potential, respectively. l stands for the position along the diffusion region of thickness d ($0 < l < d$). Based on the steady-state continuity equations and the law of mass conservation (Eq. 4), the electroneutrality assumption was

represented as Eq. 5:

$$\frac{\partial c_i}{\partial t} + \nabla \cdot N_i = 0 \quad (4)$$

$$F \sum_{i=1}^n z_i c_i = 0 \quad (5)$$

The boundary condition was set as Eq. 6. The simulation steps include a current distribution initialization and a steady step. Note that the initial defect concentration distribution of the three models is set to be the same.

$$\vec{n} \cdot J = 0 \quad (6)$$

The circuitry design of driving system

An 8-layer PCB with a 1.6 mm thick FR4 substrate was custom-designed to accommodate a 32×32 SRM-based passive crossbar array and its driving circuitry. The pulse generation and control module, powered by an STM32H743VIT6 microcontroller unit (MCU) with an ARM Cortex-M7 core operating at 480 MHz, managed timing signals and row/column addressing. Four AD5764R digital-to-analog converters (DACs) with 16-bit resolution and ± 10 V output generated programming pulses (± 2 V, 20 μ s) and read pulses (0.5 V, 20 μ s). These pulses were buffered and amplified by OPA548 operational amplifiers (Op-Amps) to drive the SRM array's rows and columns. For row/column selection, 64-channel ADG732 analog switches with ± 15 V compliance were used, reducing interconnect complexity. LT3045 low-noise voltage regulators provided stable ± 2 V and 0.5 V power supplies, while series resistors (1 k Ω) and transient voltage suppressors (TVS diodes) protected against overcurrent. Precision current-sense amplifiers (TI INA240) with a gain of 100 V/A measured sub-nanoampere currents from the SRMs, which were then digitized by ADS1256 analog-to-digital converters with 24-bit resolution and a 30 kSPS sampling rate, achieving a conductance resolution of 0.1 pS. Data communication with an external computer was facilitated by an FTDI FT601 USB 3.0 controller, supported by a Raspberry Pi Compute Module 4 for GPIO expansion. The PCB layout included guard rings and ground planes to isolate analog and digital sections, minimizing crosstalk, and featured copper pours and heatsinks for thermal management of high-power components.

ADS configurations

By combining the core architecture of the YOLO model with an FPGA-driven 32×32 SRM array, the acceleration hardware achieves real-time updating of the four layers of weights in the YOLO model, including the Data Preprocessing Layer, the Feature Extraction Layer, the Feature Fusion Layer consisting of Backbone and Neck modules, and the Head Prediction Layer. First, the data preprocessing layer extracts the local features of the image through a convolution operation (Eq. 7), followed by enhancing the feature representation using the Channel Attention Mechanism (CBAM), which weights the features through the Channel Attention Module and the Spatial Attention Module to enhance the model's attention to the important features. Finally, the spatial dimensionality of the feature map is reduced by a maximum pooling operation.

$$X_{out} = W_{conv} X_{in} + b \quad (7)$$

where W_{conv} refers to the convolution kernel weights, X_{in} refers to the input feature map, and b refers to the bias term. The feature extraction layer extracts high-level features through multiple GhostConv

modules. The GhostConv modules generate the feature maps through a combination of main convolution and cheap convolution, which significantly reduces the amount of computation (Eq. 8), where each GhostConv module is followed by a CBAM module to further enhance the feature representation.

$$\mathbf{X}_{out} = [\mathbf{Conv}_{primary}(\mathbf{X}_{in}); \mathbf{Conv}_{cheap}(\mathbf{Conv}_{primary}(\mathbf{X}_{in}))] \quad (8)$$

where $\mathbf{Conv}_{primary}$ represents the main convolutional operation and \mathbf{Conv}_{cheap} is the cheap convolutional operation, both of which are spliced through the channel. The feature fusion layer achieves multi-scale feature fusion through a bidirectional feature pyramid network (BiFPN). The BiFPN fuses features of different scales through top-down (Eq. 9) and bottom-up (Eq. 10) paths to enhance the YOLO model's detection capability for multi-scale targets.

$$\mathbf{P}_i^{td} = \mathbf{Conv}(\mathbf{P}_i + \mathbf{UpSample}(\mathbf{P}_{i+1}^{td})) \quad (9)$$

$$\mathbf{P}_i^{bu} = \mathbf{Conv}(\mathbf{P}_i^{td} + \mathbf{MaxPool}(\mathbf{P}_{i-1}^{bu})) \quad (10)$$

Finally, the dynamic detection head combines scale attention and spatial attention to weight the features for category prediction (Eq. 11) and bounding box regression (Eq. 12), respectively.

$$\mathbf{ClassPred} = \mathbf{Conv}(\mathbf{X} \odot \mathbf{ScaleAtt}(\mathbf{X})) \quad (11)$$

$$\mathbf{RegPred} = \mathbf{Conv}(\mathbf{X} \odot \mathbf{SpatialAtt}(\mathbf{X})) \quad (12)$$

where \odot denotes element-by-element multiplication.

By controlling the analog computational power of the memristor array, the FPGA achieves efficient weight programming and updating. Specifically, the FPGA communicates with the host computer through the interface, transfers the weight matrix to the SRM array in chunks using a control packet, and realizes weight quantization and programming through the DAC configuration. The FPGA divides the large weight matrix into 32×32 sub-chunks through the chunk programming strategy, and transfers and programs them into the memristor array block by block. The programming process for each sub-block consists of the following steps: first, the FPGA sends a control packet to start the programming sequence, followed by transmitting the weight data line by line, and finally sends an end packet to complete the programming. During the training phase, the model gradually updates the weights in the FPGA through a gradient accumulation strategy to ensure hardware stability. Specifically, the model calculates the loss function after each forward propagation, and calculates the gradient ∇W through back propagation with the gradient accumulation formula (Eq. 13):

$$\nabla W_{accumulated} = \sum_{i=1}^m \nabla W_i \quad (13)$$

where m is the number of gradient accumulation steps. The accumulated gradients are updated with weights through the optimizer and the updated weights are programmed into the SRM array through the FPGA interface. In the inference stage, the FPGA accelerates the forward propagation process of YOLO model through the analog computing power of the SRM array.

Subsequently, the model outputs detection results including class ID and bounding box, and maps the detection results into discrete state space through state representation function S (Eq. 14). The decision module discretizes the continuous values of the target, such as position and velocity, into a finite number of intervals, thus reducing the complexity of the state space. The discretized data are combined

into a multidimensional state vector for describing the overall state of the current environment. This provides clear state inputs to the reinforcement learning algorithm and helps the system to select optimal actions (e.g., braking, steering, accelerating, etc.). Finally, the decision-making module adopts a reinforcement learning framework for real-time decision-making based on the Q-learning algorithm. the Q-value update formula could be depicted as Eq. 15.

$$S = (\text{class ID}, \text{distance}, \text{speed}) \quad (14)$$

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha(r + \gamma \max_{a'} Q(s', a')) \quad (15)$$

where α is the learning rate, γ is the discount factor, and r is the immediate reward. The system selects the optimal policy based on the current state s and the decision action a (brake, decelerate, turn left, turn right, continue driving), and predicts the next state s' by the state transfer model. The state transfer model is estimated by counting the number of historical transfers (Eq. 16), where $C(s, a, s')$ is the number of transferring to state s' after performing action a from state s .

$$P(s', |, s, a) = \frac{C(s, a, s')}{\sum_{s''} C(s, a, s'')} \quad (16)$$

Data availability

All of the data for this work can be found in the main text and in the Supplementary Information. The training and testing figures used in this study are available in the COCO 2017 Dataset, Mapillary Traffic Sign Dataset, LISA Traffic Sign Dataset, Lyft L5 Dataset, Berkeley DeepDrive Dataset (BDD100K), which can be accessed via <https://www.kaggle.com/datasets/awsaf49/coco-2017-dataset>, <https://labelbox.com/datasets/mapillary-traffic-sign-dataset/>, <https://github.com/AminJun/lisa>, <https://paperswithcode.com/dataset/lyft-level-5-prediction>, https://www.kaggle.com/datasets/solesensei/solesensei_bdd100k, respectively. Source data are provided with this paper.

Code availability

All codes used for the hardware-based ADS configuration are available within the "Methods" section of the article and from <https://doi.org/10.5281/zenodo.15533449>⁶¹.

References

1. Wali, A., Ravichandran, H. & Das, S. A machine learning attack resilient true random number generator based on stochastic programming of atomically thin transistors. *ACS Nano* **15**, 17804–17812 (2021).
2. Zhang, G. et al. Self-rectifying memristor-based reservoir computing for real-time intrusion detection in cybersecurity. *Nano Lett.* **24**, 15707–15715 (2024).
3. Li, X. et al. Random telegraph noise in metal-oxide memristors for true random number generators: a materials study. *Adv. Funct. Mater.* **31**, 2102172 (2021).
4. Kim, S.-I. et al. Cryptographic transistor for true random number generator with low power consumption. *Sci. Adv.* **10**, eadk6042 (2024).
5. Gao, B. et al. Concealable physically unclonable function chip with a memristor array. *Sci. Adv.* **8**, 7753 (2022).
6. Shen, Z. et al. Full-vdW heterosynaptic memtransistor with the ferroelectric inserted functional layer and its neuromorphic applications. *Adv. Funct. Mater.* **35**, 2412832 (2025).
7. Lee, E.-J., Kim, J.-Y., Kim, Y.-B. & Kim, S.-K. Microwave-transparent metallic metamaterials for autonomous driving safety. *Nat. Commun.* **15**, 4516 (2024).

8. Zhou, F. & Chai, Y. Near-sensor and in-sensor computing. *Nat. Electron.* **3**, 664–671 (2020).
9. Mennel, L. et al. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62 (2020).
10. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R. & Eleftheriou, E. Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**, 529–544 (2020).
11. Chiu, Y.-C. et al. A CMOS-integrated spintronic compute-in-memory macro for secure AI edge devices. *Nat. Electron.* **6**, 534–543 (2023).
12. Wang, Y. et al. An in-memory computing architecture based on two-dimensional semiconductors for multiply-accumulate operations. *Nat. Commun.* **12**, 3347 (2021).
13. Ielmini, D. & Wong, H. S. P. In-memory computing with resistive switching devices. *Nat. Electron.* **1**, 333–343 (2018).
14. Wan, W. et al. A compute-in-memory chip based on resistive random-access memory. *Nature* **608**, 504 (2022).
15. Zhang, Y., Chu, L. & Li, W. A fully-integrated memristor chip for edge learning. *Nano-Micro Lett.* **16**, 166 (2024).
16. Zhang, W. et al. Edge learning using a fully integrated neuro-inspired memristor chip. *Science* **381**, 1205–1211 (2023).
17. Wang, C. et al. Parallel in-memory wireless computing. *Nat. Electron.* **6**, 381–389 (2023).
18. Pi, S. et al. Memristor crossbar arrays with 6-nm half-pitch and 2-nm critical dimension. *Nat. Nanotechnol.* **14**, 35 (2019).
19. Kim, K.-H. et al. A functional hybrid memristor crossbar-array/CMOS system for data storage and neuromorphic applications. *Nano Lett.* **12**, 389–395 (2012).
20. Li, Y. & Ang, K.-W. Hardware implementation of neuromorphic computing using large-scale memristor crossbar arrays. *Adv. Intell. Syst.* **3**, 2000137 (2021).
21. Fu, T., Fu, S., Wang, S. & Yao, J. Enabling reliable two-terminal memristor network by exploiting the dynamic reverse recovery in a diode selector. *Device* **2**, 100329 (2024).
22. Zhou, J., Kim, K.-H. & Lu, W. Crossbar RRAM arrays: selector device requirements during read operation. *IEEE Trans. Electron Devices* **61**, 1369–1376 (2014).
23. Zhang, H. et al. A self-rectifying synaptic memristor array with ultrahigh weight potentiation linearity for a self-organizing-map neural network. *Nano Lett.* **23**, 3107–3115 (2023).
24. Ren, S.-G. et al. Self-rectifying memristors for three-dimensional in-memory computing. *Adv. Mater.* **36**, 2307218 (2024).
25. Zhang, G. et al. Self-rectifying memristors with high rectification ratio and dynamic linearity for in-memory computing. *Appl. Phys. Lett.* **125**, 133501 (2024).
26. Li, J. et al. Sparse matrix multiplication in a record-low power self-rectifying memristor array for scientific computing. *Sci. Adv.* **9**, 7474 (2023).
27. Jeon, K. et al. Purely self-rectifying memristor-based passive crossbar array for artificial neural network accelerators. *Nat. Commun.* **15**, 129–129 (2024).
28. Kim, S.-E. et al. Sodium-doped titania self-rectifying memristors for crossbar array neuromorphic architectures. *Adv. Mater.* **34**, 2106913 (2022).
29. Ni, R. et al. Controlled majority-inverter graph logic with highly nonlinear, self-rectifying memristor. *IEEE Trans. Electron Devices* **68**, 4897–4902 (2021).
30. Lu, C. et al. Novel three-dimensional artificial neural network based on an eight-layer vertical memristor with an ultrahigh rectify ratio (>107) and an ultrahigh nonlinearity (>105) for neuromorphic computing. *Nano Lett.* **24**, 2018–2024 (2024).
31. Xu, X. et al. Fully CMOS compatible 3D vertical RRAM with self-aligned self-selective cell enabling sub-5nm scaling. In *Proc. 36th IEEE Symposium on VLSI Technology* (IEEE, 2016).
32. Kim, K. M. et al. Low-power, self-rectifying, and forming-free memristor with an asymmetric programming voltage for a high-density crossbar application. *Nano Lett.* **16**, 6724–6732 (2016).
33. Roy, K., Jaiswal, A. & Panda, P. Towards spike-based machine intelligence with neuromorphic computing. *Nature* **575**, 607–617 (2019).
34. Yang, W., Kan, H., Shen, G. & Li, Y. A network intrusion detection system with broadband WO₃-x/WO₃-x-Ag/WO₃-x optoelectronic memristor. *Adv. Funct. Mater.* **34**, 2312885 (2024).
35. Fiorillo, C. D. & Williams, J. T. Glutamate mediates an inhibitory postsynaptic potential in dopamine neurons. *Nature* **394**, 78–82 (1998).
36. Fang, Y. et al. Improvement of HfO_x-based RRAM device variation by inserting ALD TiN buffer layer. *IEEE Electron Device Lett.* **39**, 819–822 (2018).
37. Wang, K. et al. Investigation on self-rectifying properties of Pt/HfO₂/Ti with rivet-like structure based on ALD conformal technology. *Appl. Phys. Lett.* **124**, 203503 (2024).
38. Wu, T., Gao, S. & Li, Y. IGZO/WO₃-x-heterostructured artificial optoelectronic synaptic devices mimicking image segmentation and motion capture. *Small* **20**, 2309857 (2024).
39. Jamali, R., Sajjadi, M., Taherkhani, B., Abbaszadeh, D. & Moradi, A.-R. Speckle pattern analysis of PVK:rGO composite based memristor device. *Macromol. Mater. Eng.* **309**, 2400213 (2024).
40. Park, J. et al. Multi-level, forming and filament free, bulk switching trilayer RRAM for neuromorphic computing at the edge. *Nat. Commun.* **15**, 3492 (2024).
41. Kim, B., Choi, S. H., Zhu, X. Y. & Frisbie, C. D. Molecular tunnel junctions based on π -conjugated oligoacene thiols and dithiols between Ag, Au, and Pt contacts: effect of surface linking group and metal work function. *J. Am. Chem. Soc.* **133**, 19864–19877 (2011).
42. Choi, K. et al. Growth mechanism of TiN film on dielectric films and the effects on the work function. *Thin Solid Films* **486**, 141–144 (2005).
43. Jeong, H. et al. Self-supervised video processing with self-calibration on an analogue computing platform based on a selector-less memristor array. *Nat. Electron.* **8**, 168–178 (2025).
44. Liu, S. et al. Significant effects of electrode metal work function on resistive memory devices with gelatin biodielectric layer. *J. Electrochem. Soc.* **165**, G90–G95 (2018).
45. Hsu, W. et al. Electron-Selective TiO₂ Contact for Cu(In,Ga)Se₂ Solar Cells. *Sci. Rep.* **5**, 16028 (2015).
46. Ran, H. et al. Self-rectifying switching memory based on HfO_x/FeO_x semiconductor heterostructure for neuromorphic computing. *Adv. Funct. Mater.* **35**, 2418113 (2025).
47. Yadav, M. K., Mondal, A., Sharma, S. K. & Bag, A. Unveiling thermal effects on Sn-doped β -Ga₂O₃ Schottky barrier diodes on sapphire for high-temperature power electronics. *IEEE Trans. Electron Devices* **71**, 1529–1534 (2024).
48. Jang, Y. H. et al. Graph analysis with multifunctional self-rectifying memristive crossbar array. *Adv. Mater.* **35**, 2209503 (2023).
49. Cui, J. et al. CMOS-compatible electrochemical synaptic transistor arrays for deep learning accelerators. *Nat. Electron.* **6**, 292 (2023).
50. Kerber, A. & Cartier, E. A. Reliability challenges for CMOS technology qualifications with hafnium oxide/titanium nitride gate stacks. *IEEE Trans. Device Mater. Reliab.* **9**, 147–162 (2009).
51. Ren, S.-G. et al. Pt/Al₂O₃/TaO_x/Ta self-rectifying memristor with record-low operation current (<2 pA), low power (fJ), and high scalability. *IEEE Trans. Electron Devices* **69**, 838–842 (2022).
52. Sung, S. H., Kim, T. J., Shin, H., Im, T. H. & Lee, K. J. Simultaneous emulation of synaptic and intrinsic plasticity using a memristive synapse. *Nat. Commun.* **13**, 2811 (2022).
53. Sharma, D. et al. S. Linear symmetric self-selecting 14-bit kinetic molecular memristors. *Nature* **633**, 560–566 (2024).

54. Yu, F. et al. BDD100K: a diverse driving dataset for heterogeneous multitask learning. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2020).
55. Wang, C.-Y., Yeh, J., & Liao, H.-Y. M. YOLOv9: learning what you want to learn using programmable gradient information. In *Proc. 18th European Conference on Computer Vision* (Springer Nature Switzerland, 2024).
56. Qin, F., Zhang, Y., Song, H. W. & Lee, S. Enhancing memristor fundamentals through instrumental characterization and understanding reliability issues. *Mater. Adv.* **4**, 1850–1875 (2023).
57. Song, C., Liu, B., Wen, W., Li, H. & Chen Y. A quantization-aware regularized learning method in multilevel memristor-based neuromorphic computing system. In *Proc. IEEE 6th Non-Volatile Memory Systems and Applications Symposium (NVMSA)* (IEEE, 2017).
58. Yi, W. et al. Quantized conductance coincides with state instability and excess noise in tantalum oxide memristors. *Nat. Commun.* **7**, 11142 (2016).
59. D’Agostino, S. et al. DenRAM: neuromorphic dendritic architecture with RRAM for efficient temporal processing with delays. *Nat. Commun.* **15**, 3446 (2024).
60. Aguirre, F. et al. Hardware implementation of memristor-based artificial neural networks. *Nat. Commun.* **15**, 1974 (2024).
61. Zhang, G. Self-rectifying memristors with high rectification ratio for attack-resilient autonomous driving systems, Guobin-Zhang/YOLO_ADS. <https://doi.org/10.5281/zenodo.15533450> (2025).

Acknowledgements

The authors thank Jiabao Sun of ZJU Micro-Nano Fabrication Center for the support. This work is supported by the National Natural Science Foundation of China (Grants Nos. 92364204 and 62204219), the open research fund of Suzhou Laboratory (Grants No. SZLAB-1208-2024-TS012) and Major Program of Natural Science Foundation of Zhejiang Province (Grants No. LDT23F0401).

Author contributions

Yishu Zhang, Qing Wan, and Xiangshui Miao guided the research and reviewed the original manuscript. Guobin Zhang wrote the original manuscript, designed the experiment and fabricated and tested the devices. Zijian Wang, Jie Wang, Xuemeng Fan, and Bin Yu helped with the electrical measurements. Zhejia Zhang, Pengtao Li, Yitao Ma, and Kejie Huang co-designed the circuit. All authors contributed to interpreting the data and writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-60970-4>.

Correspondence and requests for materials should be addressed to Qing Wan, Xiangshui Miao or Yishu Zhang.

Peer review information *Nature Communications* thanks Min-a Kang, Guangdong Zhou and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025