

A reconfigurable photosensitive split-floating-gate memory for neuromorphic computing and nonlinear activation

Received: 27 June 2025

Accepted: 6 January 2026

Published online: 14 January 2026

 Check for updates

Zhi-Cheng Zhang¹, Yuan Li¹, Jian Yao², Zhaolong Chen³, Fu-Dong Wang¹, Shu-Han Si⁴, Yue Ding¹, Hui-Ling Qi¹, Tong-Bu Lu⁴, Lixing Kang², Zhi-Bo Liu^{1,5} & Jian-Guo Tian^{1,5} & Xu-Dong Chen^{1,6}

The rapid growth of artificial intelligence and the Internet of Things calls for compact hardware platforms that integrate sensing, computing, and nonlinear processing within a unified architecture. However, most existing neuromorphic systems implement only partial functionalities and rely on heterogeneous device integration, limiting scalability and efficiency. Here, we show a high-speed, reconfigurable multi-modal split-floating-gate memory that monolithically integrates in-sensor computing, in-memory computing, and multiple nonlinear activation functions within a single device structure. By programming charges in spatially separated floating gates, the device enables non-volatile analog control of photoresponsivity and conductance, as well as electrically reconfigurable rectification to emulate ReLU and Sigmoid activations. We further demonstrate a fully hardware-implemented sensor–processor system based on the multi-modal split-floating-gate memory arrays that performs complete unsupervised and supervised learning tasks. This work establishes a compact, energy-efficient, and reconfigurable hardware foundation for scalable intelligent systems beyond conventional silicon architectures.

The proliferation of edge devices in the era of artificial intelligence (AI) and the Internet of Things (IoT) has led to massive generation of unstructured data directly at the sensory interface^{1,2}. Conventional vision systems are typically built on a rigid architectural separation between sensing, memory, and processing modules, requiring frequent and energy-intensive data transfer from peripheral sensors to centralized processors^{3–4}. This bottleneck severely limits the real-time performance and energy efficiency of edge-intelligent systems. In contrast, biological vision systems process sensory information hierarchically and locally within neural circuits (Fig. 1a), achieving low

latency and high throughput through unified sensing–processing pathways^{5–9}.

To emulate such efficiency, recent advances in in-sensor computing (ISC)^{3,5,10–18} and in-memory computing (IMC)^{19–25} have enabled analog-domain matrix operations, significantly reducing reliance on data conversion and long-range transfer^{1,12,26,27}. While several platforms (e.g., complementary metal semiconductor oxide (CMOS) circuits^{28–36}, analog-to-digital converters (ADCs)^{37–43}, and Mott device⁴⁴) have demonstrated individual nonlinear activation functions (NAFs) in hardware, the integration of multiple, reconfigurable activation types

¹The Key Laboratory of Weak Light Nonlinear Photonics, Ministry of Education, School of Physics, Nankai University, Tianjin, China. ²Division of Advanced Materials, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China. ³School of Advanced Materials, Peking University Shenzhen Graduate School, Peking University, Shenzhen, China. ⁴MOE International Joint Laboratory of Materials Microstructure, Institute for New Energy Materials and Low Carbon Technologies, School of Material Science and Engineering, Tianjin University of Technology, Tianjin, China. ⁵State Key Laboratory of Photovoltaic Materials and Cells, Nankai University, Tianjin, China. ⁶Academy for Advanced Interdisciplinary Studies, Nankai University, Tianjin, China.

✉ e-mail: lxkang2013@sinano.ac.cn; liuzb@nankai.edu.cn; jitian@nankai.edu.cn; chenxd@nankai.edu.cn

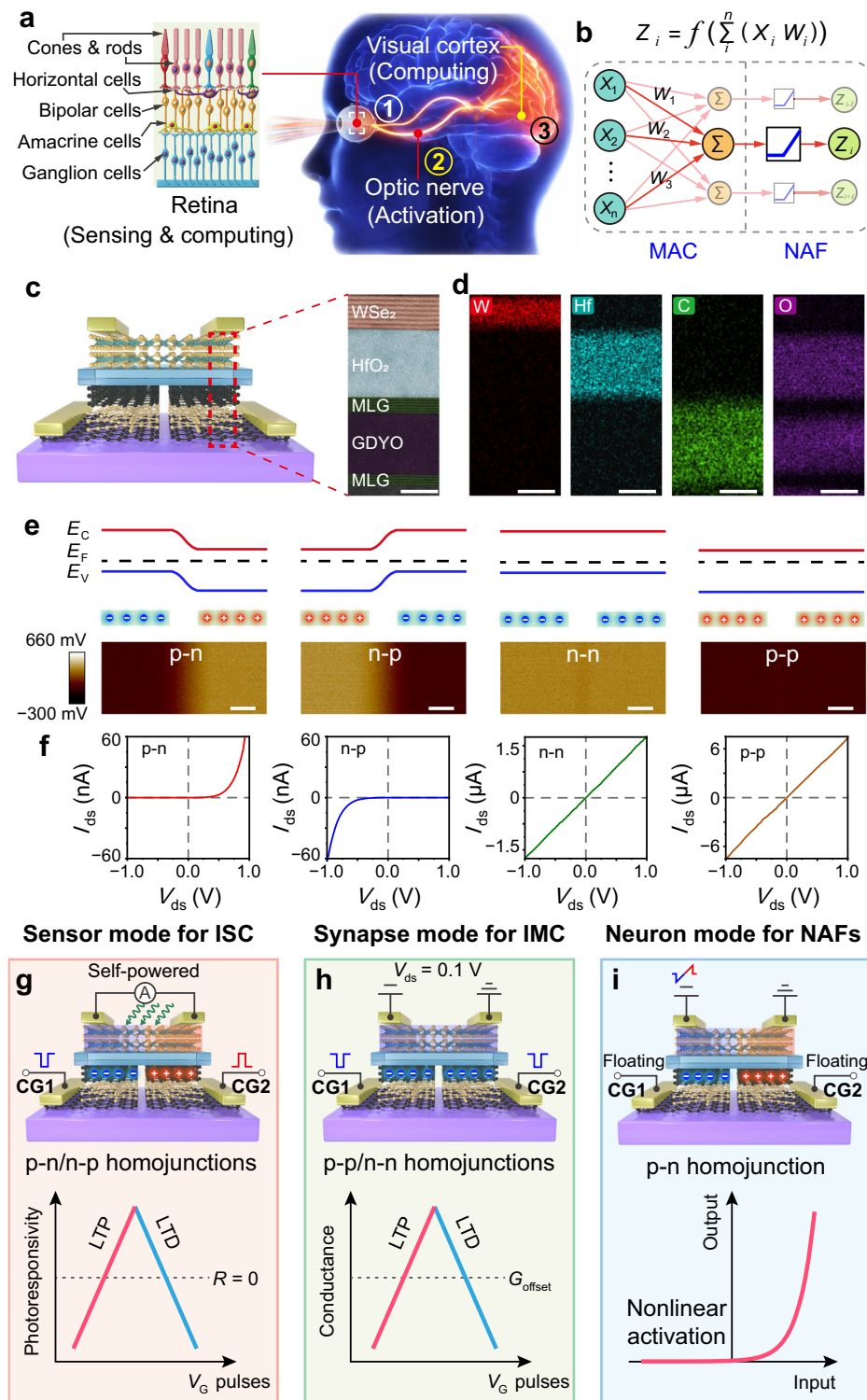


Fig. 1 | Mechanism of the MM-SFGM with reconfigurable functionalities.

a Illustration of the human visual system consisting of retina, optic nerve, and visual cortex for hierarchical processing. Visual information is first perceived and processed in the retina, where key features are extracted and significantly compressed. These compressed features are then transmitted via the optic nerve to the visual cortex for further, more complex processing in the brain. **b** Illustration of a neural network layer combining linear MAC and nonlinear activation operations in each

neuron. **c** Illustration of the structure and cross-sectional scanning transmission electron microscope (STEM) image of the MM-SFGM. Scale bar, 5 nm. **d** Energy dispersive spectrometer (EDS) images for W, Hf, C and O. Scale bars, 5 nm. Kelvin probe force microscopy (KPFM) images (**e**) and corresponding I - V curves (**f**) of the homojunction at p-n, n-p, n-n and p-p states, respectively. Scale bars, 1 μ m. Configurations of the MM-SFGM operating for ISC (**g**), IMC (**h**) and NAF (**i**), respectively.

(e.g., ReLU and Sigmoid) alongside sensing and computing capabilities within a single physical device remains an unresolved challenge. Moreover, few platforms have succeeded in implementing all core computational primitives—ISC, IMC, and NAFs—on a single material system, a limitation that hinders compactness and prevents full-stack monolithic integration.

Herein, we introduce a high-speed, reconfigurable multi-modal split-floating-gate memory (MM-SFGM) that consolidates ISC, IMC, and tunable NAFs within a unified device structure. By leveraging charge modulation in spatially decoupled floating gates, the MM-SFGM supports linear and non-volatile control over photoresponsivity and conductance, and features electrically programmable junction asymmetry for real-time switching between ReLU- and Sigmoidal-type activation modes. We further demonstrate an end-to-end, fully hardware-based sensor–processor system comprising MM-SFGM arrays that perform both supervised classification and unsupervised autoencoding tasks. Our results demonstrate a scalable platform that eliminates the traditional need to combine multiple device types for different functions, offering a compact, reconfigurable, and general-purpose solution for future sensor–processor hardware beyond neuromorphic or silicon-constrained architectures.

Results

Mechanism of the memory with reconfigurable multifunctionality

A reconfigurable MM-SFGM is developed to construct NMVS. The structure of the MM-SFGM, as shown in Fig. 1c, d and Supplementary Fig. 5, consists of an ambipolar WSe₂ channel, a HfO₂ dielectric layer, and two split multilayer graphene/graphdiyne oxide/multilayer graphene (MLG/GDYO/MLG) heterostructures. As described in Supplementary Note 1 and Supplementary Figs. 11 and 16, the GDYO layer exhibits volatile threshold-switching (TS) characteristics, which is critical for achieving nanosecond-scale programming at low voltages and long-term charge storage in the MLG floating gates⁴⁵. Under a nanosecond-scale negative or positive V_{CG} pulse, the GDYO TS layer transiently switches to a conductive state, allowing direct injection of electrons or holes into the floating gates; once the pulse ends, the TS layer spontaneously returns to the insulating state, confining the charges for long-term non-volatile modulation. In contrast, conventional flash devices rely on tunneling through thick insulating barriers, which also provide non-volatility but typically require higher voltages and longer programming times, leading to slower switching and higher energy consumption. Therefore, the adoption of GDYO is essential to enable the low-energy, high-speed, and multifunctional reconfigurability of the MM-SFGM. Details for the device fabrication is described in Methods and Supplementary Figs. 2 and 3. By controlling the type and number of charges in the split floating gates, different 2D homojunctions (including p-n/n-p/n-n/p-p) can be configured in the WSe₂ channel (Fig. 1e and f). Due to the long-term retention of charges in the floating-gates, these configurations exhibit non-volatile characteristics (Supplementary Fig. 17). Therefore, this MM-SFGM can be configured to multiple states, enabling it to perform different tasks.

Specifically, the MM-SFGM can perform ISC tasks when the homojunction is programmed to the p-n and n-p states (Fig. 1g). ISC enables the direct construction of neural networks within the sensor array, facilitating forward-propagation algorithms and convolutional operations using photoresponsivity as the weights^{3,10,12,14} (Supplementary Note 2). For ISC, a linear and non-volatile modulation of photoresponsivity (R) is required across both positive and negative regimes^{3,12,14}. The MM-SFGM operates under short-circuit condition, where self-powered positive/negative photoresponse is realized in the n-p/p-n homojunction via the photovoltaic effect (Supplementary Fig. 18). The photoresponse can be dynamically tuned from positive to negative by applying paired voltage pulses ($-V_{CG}/+V_{CG}$) to CG1 and CG2, and vice versa (Supplementary Fig. 19).

Similarly, when the MM-SFGM is configured to the p-p or n-n states, it can perform IMC tasks using the conductance (G) states as the weights (Fig. 1h and Supplementary Note 2)^{21,22,46–49}. In this case, the two control gates are connected, and identical V_{CG} pulses are applied to CG1 and CG2. The linear and non-volatile modulation of conductance is achieved by adjusting the number of restricted charges in the split floating gates, which in turn modulates the doping levels of the channel (Supplementary Fig. 26).

Moreover, the rectifying behavior of the MM-SFGM in the p-n state enables the realization of NAFs (Fig. 1i and Supplementary Note 3). By adjusting the electron and hole concentrations stored in the split floating gates, the device exhibits tunable rectification characteristics. When the electron/hole density is low, the device exhibits a ReLU-type I - V curve, while higher densities facilitate the formation of a Sigmoidal-type NAF. Consequently, the neuron-mode MM-SFGM can effectively perform NAFs on the weighted sum results in ANNs.

Sensor mode of the memory for in-sensor computing

The MM-SFGM can function as a self-powered photodiode when the homojunction is programmed to the p-n or n-p states. Figure 2a shows the spatial distribution of photocurrents generated at the p-n/n-p junction interface without drain bias. By adjusting the stored charges in the split floating-gates, the short-circuit photocurrent (I_{sc}) and photoresponsivity can be dynamically modulated between ± 80 mA W⁻¹ (Fig. 2b and Supplementary Fig. 19). Over 63 distinct states (6 bits) are accessible through long-term potentiation (LTP) and long-term depression (LTD) processes, achieved by applying paired V_{CG} pulses with opposite polarities (± 1.2 V, 20 ns). The energy required for programming is as low as 4.8 fJ. The nonlinearity, symmetry, and cycle-to-cycle variation (CCV) of the photoresponsivity update are calculated as 0.29 (LTP)/0.23 (LTD), 800, and 3.1% (LTP)/2.9% (LTD), respectively (Supplementary Note 4). These characteristics allow precise control of photoresponsivity during the update process, enabling accurate weight adjustments through the application of the corresponding number of paired V_{CG} pulses to the split control gates, as demonstrated in Fig. 2c.

The photoresponsivity modulation is non-volatile due to the long-term charge storage in the split floating-gates. As shown in Fig. 2d, I_{sc} remains stable across 17 distinct photoresponsivity states for over 1000 s. The endurance of the photoresponsivity during updates is demonstrated in Fig. 2e, where the photoresponsivity is programmed to three exemplary states (0 and ± 80 mA W⁻¹) for 10⁶ cycles, showing negligible deterioration throughout. Additionally, all 63 photoresponsivity states were programmed for 1000 receptions, showing a narrow distribution without no overlap (Fig. 2f and Supplementary Fig. 29). The photoresponse, with both positive and negative photoresponsivity, is shown in Supplementary Fig. 30a and b, with rise and fall time of 450 ns and 290 ns, respectively. Figure 2g and Supplementary Fig. 30 show the photocurrents generated by the device when triggered by optical pulses with different widths. A fully response is observed for pulse widths exceeding 500 ns, indicating the nanosecond operation speed of the MM-SFGM in sensor mode.

In addition to precise photoresponsivity update, the linear dependence of photocurrent on incident light intensity (P) is crucial for multiply-accumulate (MAC) and matrix-vector multiplication (MVM) operations in ISC hardware⁵⁰. Unlike the nonlinear power dependence of photocurrents induced by charge-trapping or redox processes^{51,52}, the I_{sc} generated by the photovoltaic effect in our device exhibits excellent linearity in a quite large power region (Fig. 2h and Supplementary Fig. 31). The analog-analog multiplication capability of the device for ISC was further validated by illuminating the device (with 7 photoresponsivities in range of ± 60 mA W⁻¹) with effective light intensities ranging from 60 to 120 nW. The generated I_{sc} matched the ideal results closely (Supplementary Fig. 32), demonstrating the high accuracy of the P - R multiplication. Furthermore, 1000 optical pulses

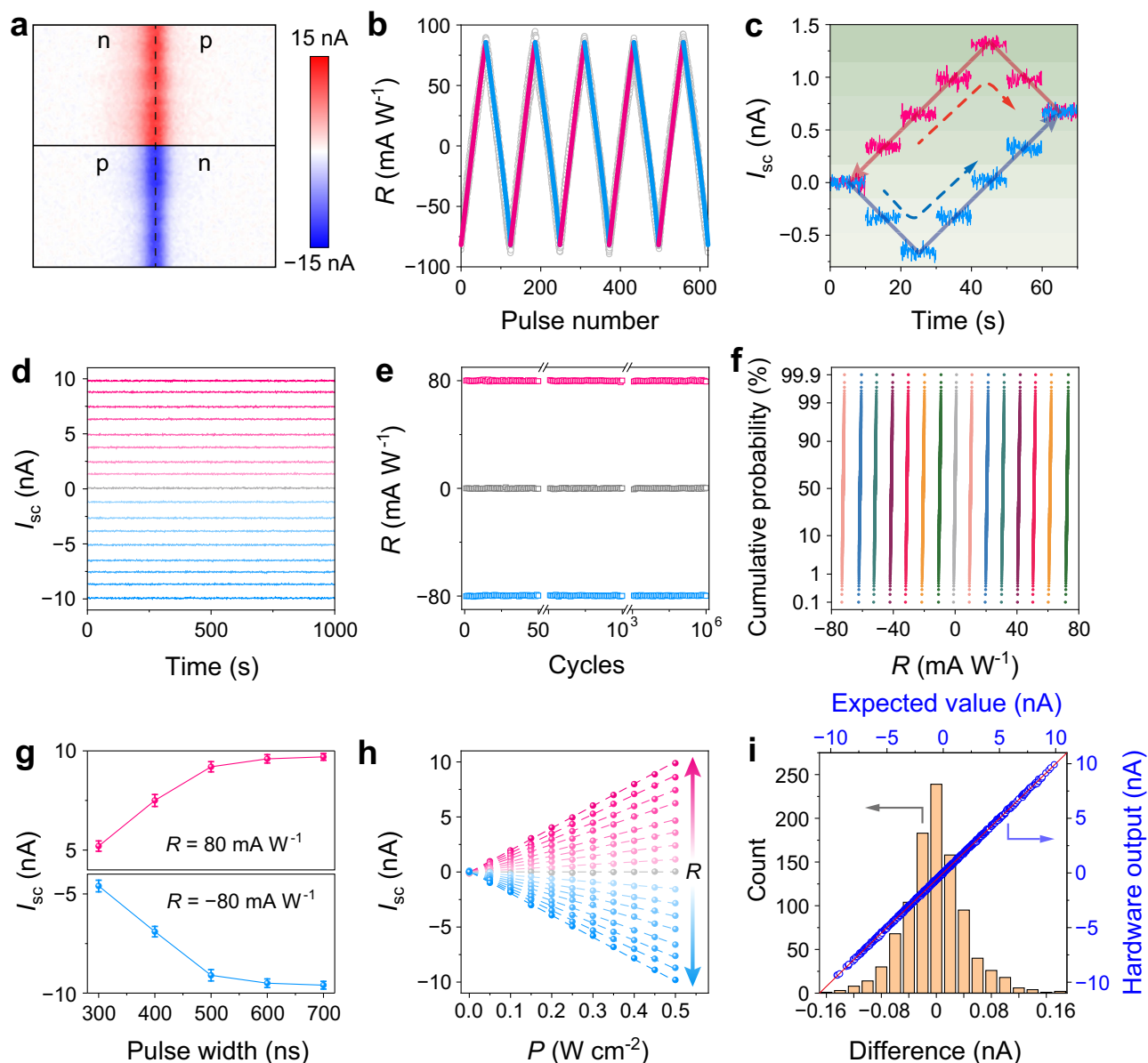


Fig. 2 | Sensor mode of the MM-SFGM for ISC. **a** Photocurrent mapping images of the homojunction at n-p and p-n states. **b** Linear and symmetric photoresponsivity update between $\pm 80 \text{ mA W}^{-1}$ by applying paired V_G pulses with opposite polarities ($\pm 1.2 \text{ V}$, 20 ns). **c** Precise modulation of the photoresponsivity between various levels by applying corresponding numbers of paired V_G pulses to the split control-gates. **d** Long-term retention characteristics of the photoresponsivity at 17 distinct states. **e** Cyclic endurance of the photoresponsivity for 10^6 cyclic test. **f** Cumulative probability distribution of the photoresponsivity with respect to 15 discrete states for 1000 repetitions. **g** Photocurrents of the device with photoresponsivity of

80 mA W^{-1} (pink) and -80 mA W^{-1} (blue) while applying an optical pulse with different widths. Error bars denote the standard deviation from 10 measurements. **h** Light-intensity dependence of the photocurrents for the device at 17 distinct photoresponsivity states. **i** Top and right axes: Linear fitting between the hardware output photocurrents and expected values of the analog-analog multiplications for 1000 randomly generated light intensities (0–125 nW) and photoresponsivities (-80 to 80 mA W^{-1}). Bottom and left axes: Statistical distribution of the current differences between experimental and ideal results.

with randomly generated intensities were applied serially to the MM-SFGM, and the hardware outputs (I_{sc}) were in good agreement with the expected results (Fig. 2i). Noteworthy, since the dark current of each photodiode under short-circuit condition is ultralow, the generated positive and negative photocurrents can accumulate directly without subtracting dark currents (Supplementary Fig. 33).

Synapse mode of the memory for in-memory computing

The IMC algorithm, capable of executing MVM operations within a memory array, is considered an ideal platform for neuromorphic computing, requiring linear and symmetric conductance updates across a wide range^{22,53,54}. As discussed in Supplementary Note 2, the

MM-SFGM can perform IMC when the entire channel is programmed between the intrinsic undoped state (low conductance) and the heavily p-doped state (high conductance). The doping level and conductance are controlled by the number of charges stored in the floating-gates (Supplementary Fig. 26). As shown in Fig. 3a and b, the device demonstrates a linear, symmetric, and non-volatile conductance update within the range of 0.1–7.0 μS . The robust endurance of conductance update is demonstrated for over 10^6 cycles (Fig. 3c), and the precise modulation of the conductance to the target levels during the update process can be realized by controlling the number of applied V_{CG} pulses (Supplementary Fig. 34), simplifying the hardware implementation of weight updating. Figure 3d and Supplementary Fig. 35

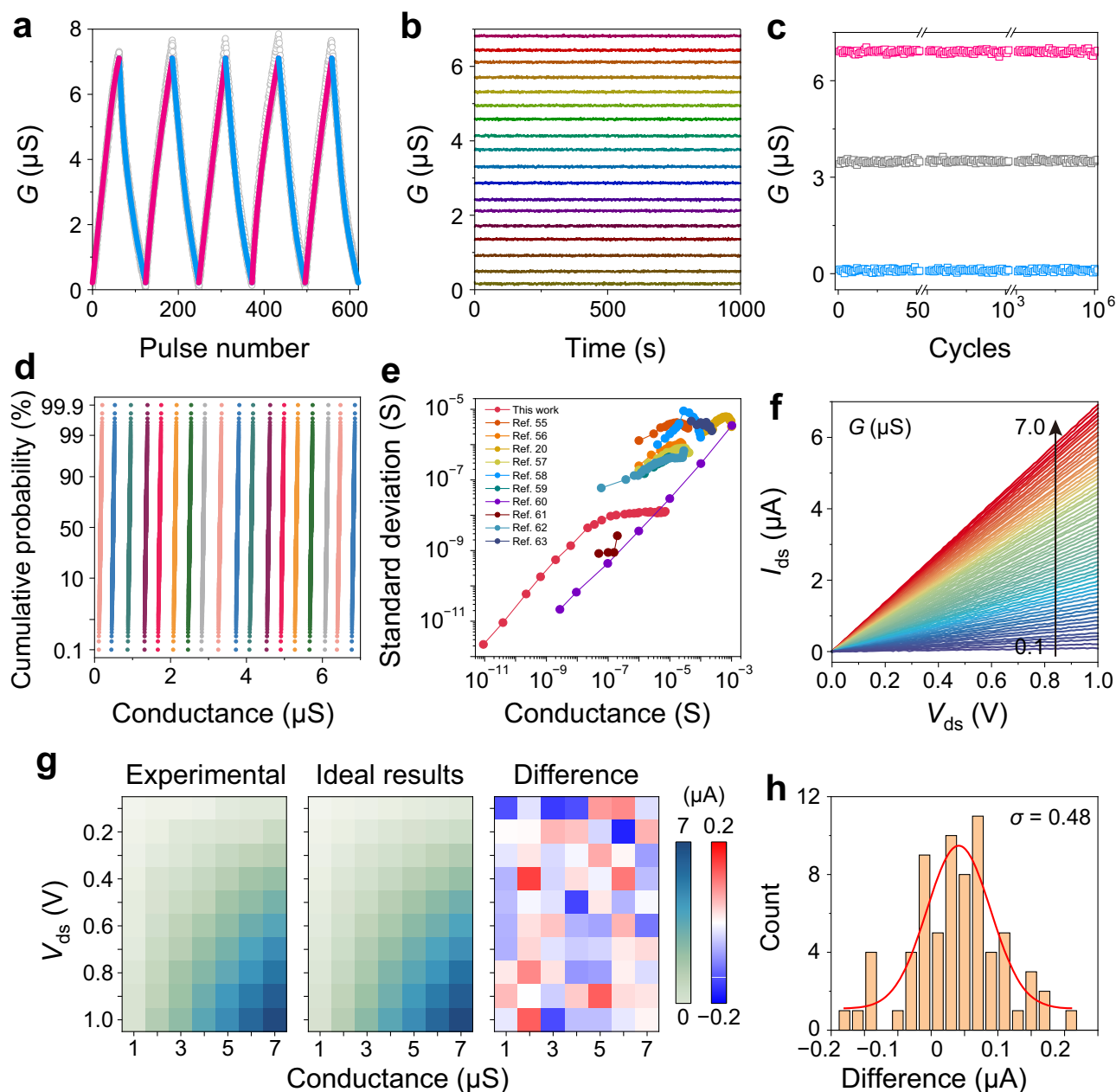


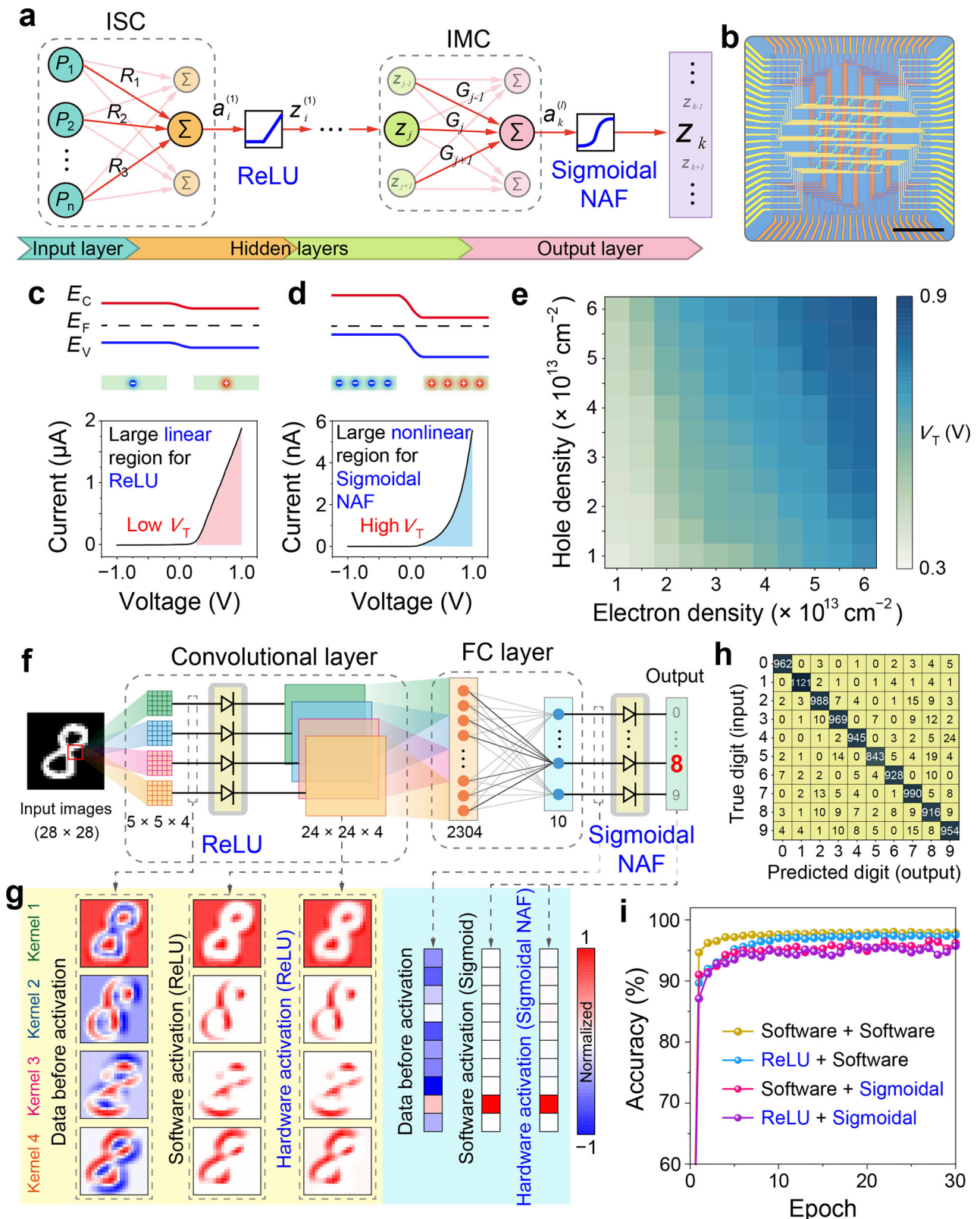
Fig. 3 | Synapse mode of the MM-SFGM for IMC. **a** Linear and symmetric conductance updating in range of 0.1–8.0 μS by applying $\pm 1.2\text{ V}/20\text{ ns}$ V_G pulses. **b** Long-term retention characteristics of the conductance at 18 distinct states. **c** Cyclic endurance of the conductance for 10^6 cyclic tests. **d** Cumulative probability distribution of the conductance with respect to 18 discrete states for 1000 repetitions. **e** Standard deviation (SD) of conductance as a function of conductance, measured over a range from 10 pS to 7 μS . Each conductance state was programmed ten times and read one hundred times after each write, with the SD

calculated from the distribution of read values. Representative data from previous reports^{20,55–59,103–106} are included for comparison. **f** Linear I - V curves of the device with 63 monotonically potentiated conductance states. **g** Analog-analog multi-multiplications between different pairs of input voltages (vertical axis) and device conductance (horizontal axis). The left and middle panels show the experimental and ideal output currents, respectively, while the right panel represents the current differences between the experimental and ideal results. **h** Statistical distribution of the current differences as presented in (g).

show the cumulative probability distribution of the MM-SFGM with 63 discrete conductance states for 1000 receptions, where all the states are well separated without overlapping. Beyond the cumulative distribution analysis within the linear region, we further evaluated the stability of 30 representative conductance states covering the entire dynamic range from 10 pS to 7 μS . Repeated programming (10 cycles) and reading (100 times per write) demonstrated standard deviations consistently below 13 nS (Fig. 3e and Supplementary Fig. 36), confirming the device's excellent stability and reliability across all operation levels. Notably, the standard deviation (SD) increases nearly linearly with conductance in the low-conductance regime

(10 pS–0.1 μS), whereas it saturates in the high-conductance regime (0.1 μS –7 μS), indicating that the device retains stable performance even at large conductance levels. To benchmark our results, Fig. 3e includes comparative data from several representative reports^{20,55–59} that exhibit narrower conductance ranges and significantly higher SD values. In contrast, the MM-SFGM maintains an ultra-wide conductance window and exceptionally low fluctuation, demonstrating excellent noise suppression and reliability compared with state-of-the-art memory and neuromorphic devices.

MVM operations inherently require a linear relationship between the output current and the input voltage. Figure 3f shows the linear I - V



relationship across all the 63 conductance states, facilitating MVM implementation. To further validate the accuracy of analog-analog multiplication for IMC, the memory was programmed into 7 equally-divided conductance (1–7 μS) and multiplied by 10 input voltages (0.1–1.0 V) in a cross-product manner. The experimental results, shown in Fig. 3g, exhibit a small deviation from the ideal results (standard deviation of 0.48 μA , Fig. 3h). Additional 1000 analog-analog

multiplication experiments were performed using randomly selected input voltages and conductance states, with the hardware outputs closely matching the expected values (Supplementary Fig. 37). The high operation speed (8 ns), as shown in Supplementary Fig. 38, enables a high-frequency computing in IMC. Similar results have been demonstrated for the device switching between the intrinsic undoped state and the heavily n-doped state (Supplementary Note 2 and

Fig. 4 | Neuron mode of the MM-SFGM for NAFs. **a** Schematic of a neural network incorporating both linear MAC operations and nonlinear activation functions. ISC and IMC hardware can be used to perform linear operations while neuron hardware can be employed to implement ReLU and Sigmoidal NAFs. **b** Optical image of a 5×5 neuron-mode MM-SFGM array for nonlinear activation. Scale bar, 100 μm . Band alignments of the slightly doped (**c**) and heavily doped (**d**) WSe₂ channel and their corresponding rectifying curves. The doping levels of the p-type and n-type regions are controlled by the concentrations of electrons and holes stored in the split floating-gates. V_T refers to the voltage at which the device transitions from nonlinear to linear I - V behavior. **e** Distribution of V_T for the device under different

electron and hole concentration conditions, extracted from Supplementary Fig. 41. **f** Schematic of a CNN consisting of one conventional layer ($5 \times 5 \times 4$) and one FC layer (2304×10). The outputs of the four conventional kernels and the output layer are processed by the ReLU-type and Sigmoidal-type hardware neurons, respectively. **g** Output results of the hardware neurons for ReLU activation in the conventional layer and Sigmoidal activation in the output layer. **h** Confusion matrix for the MNIST dataset obtained by the CNN with hardware neurons. **i** Recognition accuracies of the CNN for the MNIST dataset. Four different combinations of software and hardware implementations for ReLU and Sigmoidal NAFs were employed in the CNN.

Supplementary Fig. 39). All these characteristics highlight the great potential of the MM-SFGMs for implementing IMC.

In addition to supporting analog computing, the MM-SFGM also enables high-speed, non-volatile data storage through appropriate configuration of its conductance switching pathways (Supplementary Note 2). In particular, when the device transitions between states such as n-p and p-p (or n-n), it exhibits a high conductance modulation ratio ($\sim 10^7$), excellent retention characteristics ($>10^5$ s), robust cyclic endurance ($>10^6$ cycles), and fast switching speed (20 ns) (Supplementary Fig. 23). These characteristics reflect the intrinsic tunability of the device, allowing it to fulfill either computational or storage roles based on task-specific requirements. This versatility, achieved within a unified material and device framework, highlights its potential for constructing multifunctional, reconfigurable IMC architectures.

Neuron mode of the memory for nonlinear activations

In a complete neural network, not only linear MAC and MVM operations are required, but nonlinear activation operations are also essential. Typically, the hidden layers in a neural network use ReLU activation, while the output layer utilizes NAFs like Sigmoid or Softmax (Fig. 4a). By exploiting the tunable rectification characteristics of the MM-SFGM in the p-n state, it can simulate both ReLU and Sigmoidal NAFs. As shown in Supplementary Fig. 40, the impact of electron and hole densities in the split floating-gates on the rectifying behavior of the device was systematically studied. As discussed in Supplementary Note 3, when the concentrations of electrons and holes in the floating-gates are low, the built-in potential of p-n junction is reduced, and the rectified I - V curve shows a large linear region (Fig. 4c), which is advantageous for ReLU operation. In contrast, when the electron and hole concentrations are high, the built-in potential increases, and the nonlinear region dominates in the I - V curve (Fig. 4d), making it suitable for Sigmoidal activation. Figure 4e shows the distribution of V_T for the device under different electron and hole concentration conditions, where V_T represents the voltage at which the device transitions from nonlinear to linear I - V behavior. A smaller V_T results in a larger linear region, making it more suitable for ReLU operations. Conversely, a larger V_T leads to a larger nonlinear region, making it more suitable for Sigmoidal NAF operations. As shown in Supplementary Fig. 45, the device also demonstrated nanosecond-scale operation speeds in neuron mode, performing both ReLU and Sigmoidal NAF operations with high efficiency.

To verify the ability of our device to perform nonlinear activation in neural networks, a convolutional neural network (CNN), as illustrated in Fig. 4f, was constructed. In this CNN, the output data from both the convolutional layer and the fully-connected (FC) layer were nonlinearly activated by the neuron-mode MM-SFGM hardware, as shown in Fig. 4b and Supplementary Fig. 46. The consistency of the devices' performance for both ReLU and Sigmoidal operations is demonstrated in Supplementary Figs. 66 and 67. Figure 4g and Supplementary Fig. 48 show the output results of the hardware neurons for ReLU activation in the conventional layer and Sigmoidal activation in the output layer, with performance closely matching the software simulation results. Figure 4h and i display the recognition results of the CNN using hardware neurons for the MNIST dataset, achieving a high

accuracy of 95.2%, which is comparable to the accuracy of the fully software-simulated model (97.9%).

Reconfigurable device arrays for neuromorphic machine vision

The reconfigurable multifunctionalities of the MM-SFGM enable the construction of a NMVS integrating visual information sensing, computing, and NAFs. To construct a fully hardware neural networks involving ISC, IMC, and NAFs, here three $5 \times 5 \times 4$ device arrays were fabricated to perform ISC, IMC, and NAFs (Fig. 5a and Supplementary Figs. 20 and 46). The performance consistency of these as-fabricated devices is demonstrated in Supplementary Note 4. Figure 5b and illustrates the structure of a fully hardware NMVS based on the MM-SFGM arrays (details in Supplementary Note 5). In this system, sensor-mode and synapse-mode device arrays are used to constitute 25×4 and 4×25 ANNs to implement ISC and IMC, respectively. Additionally, neuron-mode devices are employed to perform ReLU and Sigmoidal NAFs. This NMVS enables both supervised and unsupervised tasks for visual information.

As a demonstration, classification and autoencoding tasks were performed by this NMVS for a custom dataset containing four types of letters ("F", "L", "R" and "S", 5×5 pixels) with Gaussian noises ($\sigma = 0.3$) (Supplementary Fig. 76). Optical images were projected into the ISC hardware via the equipment as shown in Supplementary Fig. 77. For classification, the four output photocurrents from the ISC device array (Supplementary Fig. 78) are nonlinearly activated using Sigmoidal-type NAF devices (Supplementary Fig. 79). The four activated results as shown in Fig. 5c represent the probabilities of letters "F", "L", "R", and "S", respectively, where the substantially higher value of the target letter indicates the successful identification. Details for the training of the ANN is described in Methods. Supplementary Fig. 80 depict the photoresponsivity distribution of the ANN before and after 30 epochs, and the evolution of the recognition accuracy and loss during training are shown in Fig. 5d.

For the unsupervised autoencoding task, ISC array functions as the encoder and IMC array acts as the decoder. As shown in Fig. 5e, the four output photocurrents from ISC array are activated by the ReLU-type NAF devices, as the encoded data fed into the IMC array (Supplementary Fig. 72). The 25 output current signals from the IMC array (Supplementary Fig. 81) are further nonlinearly activated by the Sigmoidal NAF devices and are reconstructed into a 5×5 -pixel image. The training process is described in Methods and Supplementary Note 5, and the weights distributions (photoresponsivity in ISC and conductance in IMC) of the NMVS before and after training are shown in Supplementary Fig. 82. Supplementary Fig. 83 depicts the reconstructed images of the trained NMVS for various noisy input images, and the output images of the NMVS remained almost unchanged for over 12 h owing to the local storage of the weights in the networks after training (Supplementary Figs. 84 and 85).

Notably, due to the non-volatile modulation of the photoresponsivity and conductance, the weights can be directly stored in the NMVS hardware after training, without any additional energy consumption. This contrasts with the volatile-device-based ISC/IMC hardware, which requires external power to maintain the weights. Specifically, for the ISC hardware, the sensor-mode

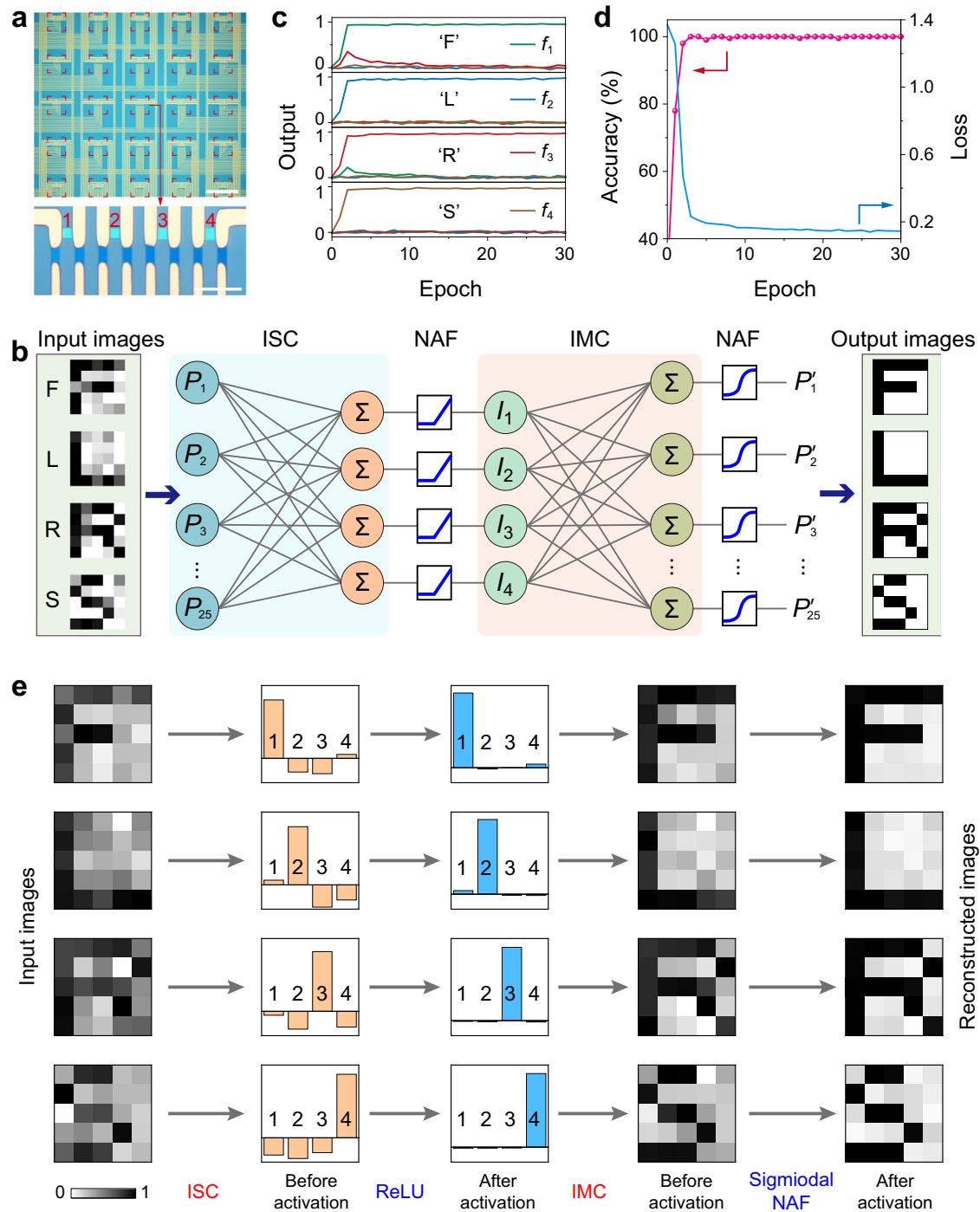


Fig. 5 | Reconfigurable MM-SFGM arrays for NMVS. a Optical images of a $5 \times 5 \times 4$ MM-SFGM array. Scale bars, 300 μm and 20 μm . **b** Schematic of a NMVS consisting of a 25×4 ISC device array, a 4×25 IMC array, and ReLU/Sigmoidal-type NAF devices. **c** Average output signals (f_1, f_2, f_3 , and f_4) for letters “F”, “L”, “R”, and “S” as a function of training epochs for classification task. **d** Evolution of the classification accuracies and loss of the NMVS with training epochs. **e** Autoencoding process of

5×5 -pixel images with Gaussian noises ($\sigma = 0.3$). The encoding process is performed by the sensor-mode device array via ISC, outputting four photocurrent signals and being activated by the ReLU-type NAF devices. The decoding process is executed by the synapse-mode device array via IMC, with the output 25 current signals being activated by the Sigmoidal-type NAF devices, thereby reconstructing 5×5 -pixel images.

devices, which operate based on the photovoltaic effect in short-circuit current condition, sense and process optical signals in a self-powered manner. This passive current generation mechanism does not require an external power source. However, it is important to note that while the sensor array can generate output photocurrent without power consumption, the subsequent processing of the signal, such as amplification, analysis, and conversion, does require power.

Discussion

Table 1 benchmarks the MM-SFGM against representative reconfigurable neuromorphic platforms. Most previously reported “all-in-one” devices integrate only subsets of neuromorphic primitives (e.g., ISC and IMC), or demonstrate nonlinear activation as an isolated function. In contrast, the MM-SFGM unifies ISC, IMC, and reconfigurable nonlinear activation within a single device, enabling all essential neuromorphic operations to be executed on the same physical hardware.

Table 1 | Key parameters and functionalities of reconfigurable devices

Structure	Reconfigurable functions				Sensor mode				Neuron mode				Endurance	Scalability	Refs.				
	Synapse mode		Operation Speed		V _{programming}		E		Operation Speed		Mode					Operation Speed		E	
	V _{programming}	E	Operation Speed	V _{programming}	V _{programming}	E	Operation Speed	Mode	Operation Speed	Mode	Operation Speed	E							
MoTe ₂	V _g -5 V, V _d ±20 V	9.5 fJ	N/A	-	-	-	-	-	-	-	-	-	500	-	-	67			
ReSe ₂ /hBN	-40 V, 0.5 s	40 V, 1.5 s	N/A	-	-	-	-	-	-	-	-	-	-	-	-	68			
Bi ₄ Ti ₃ O ₁₂ /MoS ₂ /Graphene	±10 V	300 ms	N/A	-	-	-	-	-	-	-	-	-	5 × 10 ⁵	-	-	69			
MoS ₂ /HfO ₂ /LiNbO ₃	-1.5 V	500 μs	N/A	-	-	-	-	-	-	-	-	-	40	3 × 3	-	70			
α-In ₂ Se ₃	±4 V, 0.2 s	N/A	N/A	-	-	-	-	-	-	-	-	-	1000	-	-	71			
TiO ₂	±4 V, 42 ns	0.147 fJ	N/A	-	-	-	-	-	-	-	-	-	10 ⁴	-	-	72			
MoS ₂ /hBN/Gr	±40 V, 1 ms	N/A	N/A	-	-	-	-	-	-	-	-	-	300	-	-	73			
BP/Al ₂ O ₃ /HfO ₂ /Al ₂ O ₃	-4 V-12 V	20 ms	N/A	±18 V, 20 ms	-	-	-	N/A	N/A	-	-	-	200	3 × 4	-	74			
PZT/SRO/STO	±3 V, 0.15 ms	N/A	2.6 μs	-1.94 V, 10 μs, +1.89 V, 10 μs	-	-	-	3.1 nJ	1 ns	-	-	-	10 ⁶	3 × 3	-	15			
MoS ₂	±10 V, 10 s	N/A	N/A	±10 V, 10 s	-	-	-	N/A	N/A	-	-	-	-	3 × 3	-	75			
WSe ₂ /Al ₂ O ₃ /HfO ₂	-6 V, 5 ms	+4 V, 5 ms	N/A	±6 V, 5 μs	-	-	-	N/A	5 μs	-	-	-	-	3 × 3	-	13			
CeO ₂ /STO	±10 V, 0.5 s	N/A	N/A	±10 V, 0.5 s	-	-	-	N/A	12 ns	-	-	-	-	3 × 3	-	76			
PZT/SRO	±3.5 V, 0.5 μs	1.77 nJ	10 μs	±10 V, 100 ns	-	-	-	56 fJ	30 μs	-	-	-	10 ⁹	2 × 2	-	77			
Gr/3R-WS ₂ /Gr	±2 V, 0.5 ms	N/A	N/A	±0.05 V, 1 s	-	-	-	N/A	19.2 μs	-	-	-	2000	3 × 3 × 3	-	78			
SWNT@GDY	±0.5 V, 1 ms	50 aJ	N/A	±0.5 V, 1 ms	-	-	-	50 aJ	5 ms	-	-	-	10 ⁵	3 × 3 × 3	-	12			
Gr/C/IPS/Gr	±1.5 V, 0.1 s	N/A	N/A	±0.5 V, 0.1 s	-	-	-	N/A	N/A	-	-	-	10 ⁵	10 × 10	-	79			
MoTe ₂ /P(VDF-TrFE)	±39 V, 50 μs	N/A	N/A	±25 V, 10 μs	-	-	-	N/A	1.9 μs	-	-	-	10 ⁶	3 × 3 × 3	-	3			
PEDOT:PSS/perovskite/C60/bathocuproine	±4 V, 10 s	N/A	N/A	±4 V, 10 s	-	-	-	3.63 fJ	2.75 ns	-	-	-	-	64 × 64	-	8			
TiO ₂ /ZnO	±3.5 V, 100 ns	2 nJ	N/A	±3.5 V, 100 ns	-	-	-	2 nJ	N/A	-	-	-	10 ⁹	90 × 8	-	80			
MoS ₂ /BN/Se	±30 V, 10 ms	N/A	N/A	±30 V, 10 ms	-	-	-	N/A	N/A	-	-	-	3000	3 × 3	-	81			
VO ₂	3 V, 1 μs	3.9 nJ	30 ns	-	-	-	-	-	-	-	-	-	10 ¹²	15 × 15	-	82			
HfZrO	+3.8 V, 100 μs -2 V, 100 μs	N/A	N/A	-	-	-	-	-	-	-	-	-	2 × 10 ⁷	-	-	83			
Hf _{0.5} Zr _{0.5} O ₂ /SiO ₂	+5 V, 500 ps -3.5 V, 500 ps	0.12 fJ	N/A	-	-	-	-	-	-	-	-	-	10 ⁷	-	-	84			
Si/HfSe ₂	±1 V, 1 ns	N/A	60 ns	-	-	-	-	-	-	-	-	-	2.65 × 10 ⁴	32 × 32	-	24			
TiO _x	4.5 V, 10 μs	N/A	100 μs	-	-	-	-	-	-	-	-	-	5 × 10 ⁶	20 × 20	-	85			
BP/Al ₂ O ₃ /WSe ₂ /h-BN	±14 V	N/A	N/A	-6 V + 5 V Persistent voltage pulse	-	-	-	N/A	100 μs	5 ms	-	-	55	3 × 3	-	4			
VO ₂	-	-	-	-	-	-	-	-	-	-	-	-	65 ns	199.5 pJ	5000	44			
μ-Si multimodal transistor	-	-	-	-	-	-	-	-	-	-	-	-	N/A	N/A	-	86			
TaO _x memristor circuit	-	-	-	-	-	-	-	-	-	-	-	-	N/A	N/A	-	87			
p(V3D3-co-VI)/Al ₂ O ₃ neural circuit	-	-	-	-	-	-	-	-	-	-	-	-	N/A	0.5 pJ	-	88			
NbO ₂ neural circuit	-	-	-	-	-	-	-	-	-	-	-	-	500 μs	1.65 pJ	5000	89			

Table 2 | Comparison of key parameters of nonlinear activation devices and circuits

Function	Type	Energy (power) consumption	Area	Delay time	Refs.
Sigmoid, Hyperbolic tangent, Soft-sign, Softplus, ELU, SeLU	CMOS	3.7 mW (5 bit), 4.6 mW (8 bit)	0.003 mm ² @16 nm	N/A	28
Tanh	CMOS	0.071 mW	582 μm ² @90 nm	1.40 ns	29
ReLU, Leaky ReLU, Tanh	CMOS	2.155 mW	35.809 μm ² @180 nm	420 ps	30
Sigmoid	CMOS	11.4 μW	184 μm ² @180 nm	N/A	31
Tanh, Sigmoid	CMOS	0.07 mW	1554 μm ² @180 nm	2.25 ns	32
ReLU, Sigmoid	CMOS	55–88 mW	0.028 mm ² @45 nm	570 ps	35
ReLU	Mott	199.5 pJ	0.64 μm ²	61.4 ns	44
ReLU	ADC	29.6 fJ	45 μm ²	N/A	37
ReLU, Tanh, Sigmoid	ADC	4.75 nJ	0.086 mm ² @90 nm	207 ns	38
ReLU	ADC	87 μW	7200 μm ² @180 nm	2.1 μs	39
Tanh, clipped-ReLU	ADC	15 μW	101.64 μm ² @28 nm	N/A	41
ReLU	ADC	3.08 TOPS/W	3.91 TOPS/mm ² @28 nm	N/A	42
Mish, GELU, SiLU	ADC	168 μW	1218 μm ² @65 nm	N/A	43
Tanh, ReLU	Digital processors	1982 mJ	263 mm ²	N/A	94
Tanh	ASIC	N/A	17864.24 μm ² @180 nm	2.45 ns	95
Sigmoid Tanh	ASIC	2.16 mW, 2.64 mW	7973.2 μm ² @65 nm, 10180 μm ² @65 nm	3 ns, 3.16 ns	96
Sigmoid	ASIC	2.66 mW	8097.48 μm ² @90 nm	1.96 ns	97
ReLU	ASIC	20.4 mW	130 nm	14 ns	98
Sigmoid	FPGA	19 mW	N/A	55 ns	99
Tanh	FPGA	N/A	N/A	45.4 ns	100
Sigmoid, Tanh	ASIC FPGA	0.31 mW, 0.28 mW, 56 mW, 56 mW	4678.78 μm ² @180 nm, 5582.5 μm ²	1.6 ns, 1.7 ns, 23.1 ns, 23 ns	101
Sigmoid	MRAM	18.04 μW	0.138 μm ²	48 ns	102
ReLU, Sigmoid	Memory	16 fJ for ReLU, 80 aJ for Sigmoid	15 μm ²	<8 ns	This work

@* nm" represents node process technology.

N/A Not available.

(FIB) with a gap of approximately 110 nm (Supplementary Fig. 4). The top MLG layer was pretreated with UV-ozone for 10 s to introduce surface dangling bonds, which facilitated the deposition of a 10.5 nm HfO₂ dielectric layer on the heterostructure via atomic layer deposition (ALD, Supplementary Fig. 6). Subsequently, CVD-grown multilayer WSe₂ (~5.6 nm, Supplementary Fig. 53) channels were stacked on HfO₂, and the source and drain electrodes (Cr/Au 10/50 nm) were deposited. Finally, the device array was mounted and wire-bonded to a probe-card with 400 pins (Supplementary Fig. 77a) for measurement. The CVD growth was performed using Micro-STSI200 system (Units Technology, www.units-tech.com.cn). Supplementary Fig. 3 presents the corresponding atomic force microscope (AFM) images during fabrication, highlighting the clean surface of each layer.

Measurement setup

The basic electrical characteristics of the devices were measured by a semiconductor analyzer (Keithley 4200A-SCS) equipped with a pulse measurement unit (PMU). An arbitrary waveform generator (Tektronix, AFG31052) was used to generate nanosecond gate pulses for the reconfiguration of the devices and updating of the weights. The photocurrent maps generated by the p-n and n-p homojunctions (Fig. 2a) were measured using a WITec (Alpha 300RAS) photocurrent testing module. Prior to testing, the MM-SFGM was programmed to the p-n or n-p states and placed on the testing platform with both CG1 and CG2 floating. A 532 nm laser with an intensity of 100 nW was focused through a 100× objective lens onto the WSe₂ channel. The source and drain electrodes were connected to a transimpedance amplifier (TIA, DHPCA-100) to measure the photocurrents generated at each position. The

photocurrent map (100 × 100 pixels) was generated by moving the device with a piezoelectric platform, with movement and measurement controlled automatically by the WITec system. For the ISC, 5 × 5-pixel images were generated by a 532 nm laser and a spatial light modulator (SLM) operating in intensity-modulation mode. Then the input images were adjusted to 1 mm × 1 mm via optical lenses, and projected to the 5 × 5 × 4 sensor array (Supplementary Fig. 77).

Implementation of image classification via NMVS

As illustrated in Supplementary Fig. 79, the NMVS for classification task consists of a 5 × 5 × 4 sensor-mode device array for ISC and four neuron-mode devices for NAF. The 5 × 5 × 4 device array is divided into 5 × 5 pixels and each pixel consists of 4 subpixels. ISC operations were executed by connecting the devices with the same subpixel index ($m = 1, 2, 3, 4$) in each pixel in parallel and accumulating the generated photocurrents (Supplementary Fig. 78). The initial weight distribution (photoresponsivity) of the sensor array was randomly generated. For each training epoch, 5 × 5-pixel images from the training dataset were optically projected to the sensor array, and the parallelly connected circuit outputted four photocurrent signals I_m :

$$I_m = \sum_{i=1}^5 \sum_{j=1}^5 I_{m,ij} = \sum_{i=1}^5 \sum_{j=1}^5 P_{ij} \cdot R_{m,ij}, \quad (1)$$

where i and j represent the row and column index. I_m was further activated by four Sigmoidal-type NAF devices, outputting four signals (f_1, f_2, f_3 , and f_4) representing the probabilities for letters "F", "L", "R",

and “S”. After each epoch, the gradients of loss function L (here cross-entropy function) was backpropagated to update the weights of the neural network:

$$\Delta R_{m,ij} = \eta \sum_{k=1}^S \nabla_{R_{m,ij}} L_k, \quad (2)$$

$$L = - \sum_{k=1}^M y_k \log \hat{y}_k, \quad (3)$$

where y_k and \hat{y}_k were the label and prediction, respectively, and learning rate $\eta = 0.2$ and the size of batch $S = 20$. Backpropagation was executed by the FPGA integrated on the same PCB (Supplementary Fig. 71). Based on computed gradients, the FPGA and MUX array coordinated voltage pulses to selectively update the photoresponsivity of sensor-mode devices after each epoch, enabling efficient in-situ learning (Supplementary Note 5.1). Classification accuracy was evaluated by projecting test images during each epoch.

Implementation of image autoencoding via NMVS

The structure of the NMVS for autoencoding task is shown in Fig. 5b and Supplementary Figs. 71 and 72, which consists of a 25×4 ANN for ISC, a 4×25 ANN for IMC, 4 ReLU-type and 25 Sigmoidal-type NAF devices. The initial weights (photoresponsivity for ISC and conductance for IMC) of the ANNs were randomly generated. For each training epoch, 5×5 -pixel noisy images chosen from the training dataset were optically projected onto the sensor array, outputting four photocurrents I_m . A ReLU operation was performed in the following via the four ReLU-type NAF devices, and the output current signals were converted to voltage and fed into the 4×25 IMC array (Supplementary Fig. 81). The 25 output signals were further nonlinearly activated by the 25 Sigmoidal-type NAF devices, which were used to reconstruct 5×5 -pixel images. The loss function (mean-square error, $L = \sum_{i=1}^5 \sum_{j=1}^5 (p_{i,j} - p'_{i,j})^2$, where $P_{i,j}$ and $P'_{i,j}$ correspond to the intensity of the (i, j) pixel in the input and reconstructed images, respectively) was backpropagated via the on-board FPGA to compute gradients. Based on these results, the FPGA and MUX array coordinated voltage pulses to update the photoresponsivity and conductance of ISC and IMC devices, respectively, after each training epoch (Supplementary Note 5.1).

Data availability

The data generated in this study have been deposited in Figshare at https://figshare.com/articles/dataset/Source_data_for_NCOMMMS-25-48768-T/30325393.

References

- Chai, Y. In-sensor computing for machine vision. *Nature* **579**, 32–33 (2020).
- Zhou, F. & Chai, Y. Near-sensor and in-sensor computing. *Nat. Electron.* **3**, 664–671 (2020).
- Wu, G. et al. Ferroelectric-defined reconfigurable homojunctions for in-memory sensing and computing. *Nat. Mater.* **22**, 1499–1506 (2023).
- Zhang, Z. et al. All-in-one two-dimensional retinomorphic hardware device for motion detection and recognition. *Nat. Nanotechnol.* **17**, 27–32 (2022).
- Wang, C.-Y. et al. Gate-tunable van der Waals heterostructure for reconfigurable neural network vision sensor. *Sci. Adv.* **6**, eaba6173 (2020).
- Wang, S. et al. An organic electrochemical transistor for multi-modal sensing, memory and processing. *Nat. Electron.* **6**, 281–291 (2023).
- Hou, Y.-X. et al. Large-scale and flexible optical synapses for neuromorphic computing and integrated visible information sensing memory processing. *ACS Nano* **15**, 1497–1508 (2021).
- He, Z. et al. Perovskite retinomorphic image sensor for embodied intelligent vision. *Sci. Adv.* **11**, eads2834 (2025).
- Zhang, Z.-C. et al. Synthesis of wafer-scale graphdiyne/graphene heterostructure for scalable neuromorphic computing and artificial visual systems. *Nano Res.* **14**, 4591–4600 (2021).
- Mennel, L. et al. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62–66 (2020).
- Yang, Y. et al. In-sensor dynamic computing for intelligent machine vision. *Nat. Electron.* **7**, 225–233 (2024).
- Zhang, G.-X. et al. Broadband sensory networks with locally stored responsivities for neuromorphic machine vision. *Sci. Adv.* **9**, eadi5104 (2023).
- Zhou, Y. et al. Computational event-driven vision sensors for in-sensor spiking neural networks. *Nat. Electron.* **6**, 870–878 (2023).
- Yan, X. et al. Reconfigurable mixed-kernel heterojunction transistors for personalized support vector machine classification. *Nat. Electron.* **6**, 862–869 (2023).
- Cui, B. et al. Ferroelectric photosensor network: an advanced hardware solution to real-time machine vision. *Nat. Commun.* **13**, 1707 (2022).
- Lee, J.-J. et al. Polarization-sensitive in-sensor computing in chiral organic integrated 2D pn heterostructures for mixed-multimodal image processing. *Nat. Commun.* **16**, 1–11 (2025).
- Cui, K. et al. Spectral convolutional neural network chip for in-sensor edge computing of incoherent natural light. *Nat. Commun.* **16**, 81 (2025).
- Xiong, Z. et al. Parallelizing analog in-sensor visual processing with arrays of gate-tunable silicon photodetectors. *Nat. Commun.* **16**, 1–13 (2025).
- Huang, X. et al. An ultrafast bipolar flash memory for self-activated in-memory computing. *Nat. Nanotechnol.* **18**, 486–492 (2023).
- Song, W. et al. Programming memristor arrays with arbitrarily high precision for analog computing. *Science* **383**, 903–910 (2024).
- Rao, M. et al. Thousands of conductance levels in memristors integrated on CMOS. *Nature* **615**, 823–829 (2023).
- Yao, P. et al. Fully hardware-implemented memristor convolutional neural network. *Nature* **577**, 641–646 (2020).
- Yue, W. et al. Physical unclonable in-memory computing for simultaneous protecting private data and deep learning models. *Nat. Commun.* **16**, 1031 (2025).
- Jain, S. et al. Heterogeneous integration of 2D memristor arrays and silicon selectors for compute-in-memory hardware in convolutional neural networks. *Nat. Commun.* **16**, 2719 (2025).
- Ortner, T. et al. Rapid learning with phase-change memory-based in-memory computing through learning-to-learn. *Nat. Commun.* **16**, 1243 (2025).
- Wang, S. et al. Networking retinomorphic sensor with memristive crossbar for brain-inspired visual perception. *Natl. Sci. Rev.* **8**, nwaa172 (2021).
- Shi, L. et al. Flexible retinomorphic vision sensors with scotopic and photopic adaptation for a fully flexible neuromorphic machine vision system. *SmartMat* **5**, e1285 (2024).
- Yang, J. et al. Efficient nonlinear function approximation in analog resistive crossbars for recurrent neural networks. *Nat. Commun.* **16**, 1136 (2025).
- Liu, K., Shi, W., Huang, C. & Zeng, D. Cost effective Tanh activation function circuits based on fast piecewise linear logic. *Microelectron. J.* **138**, 105821 (2023).
- Krestinskaya, O., Choubey, B. & James, A. Memristive GAN in analog. *Sci. Rep.* **10**, 5838 (2020).
- Krestinskaya, O., Salama, K. N. & James, A. P. Learning in memristive neural network architectures using analog back-propagation circuits. *IEEE Trans. Circuits Syst. I Regul. Pap.* **66**, 719–732 (2018).

32. Nguyen, V.-T. et al. An accurate and compact hyperbolic tangent and sigmoid computation based stochastic logic. In *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 386–390 (IEEE, 2021).
33. Priyanka, P., Nisarga, G. & Raghuram, S. CMOS implementations of rectified linear activation function. In *International Symposium on VLSI Design and Test*, **892**, 121–129 (Springer Singapore, 2019).
34. Chunarkar, S. & Chiluveru, S. R. Efficient hardware implementation of nonlinear activation function for inference model. In *2024 21st International SoC Design Conference (ISOCC)*, 350–351 (IEEE, 2024).
35. Shoobi, A. et al. Integrated photonic-electronic deep neural networks: from sub-nanosecond image classification to PVT-tolerant activation functions. In *2025 IEEE Custom Integrated Circuits Conference (CICC)*, 1–6 (IEEE, 2025).
36. Mouny, P.-A. et al. Towards a cryogenic CMOS-memristor neural decoder for quantum error correction. In *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, **1**, 1258–1263 (IEEE, 2024).
37. Im, J. et al. Multifunctional in-memory analog-to-digital converter for next-gen compute-in-memory systems. *Adv. Intell. Syst.* **7**, 2400594 (2024).
38. Giordano, M. et al. Analog-to-digital conversion with reconfigurable function mapping for neural networks activation function acceleration. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **9**, 367–376 (2019).
39. Shi, A. et al. Specific ADC of NVM-based computation-in-memory for deep neural networks. *IEEE Trans. Circ. Syst. I Regul. Pap.* **71**, 5387–5399 (2024).
40. Xie, C., Shao, Z., Zhang, M., Du, Y., Du, L. RAC-NAF: a reconfigurable analog circuitry for nonlinear activation function computation in computing-in-memory. In *IEEE J. Solid-State Circuits*, 1–11 (IEEE, 2025).
41. Lalen, N., Padma, S., Kämpfe, T., Jang, T. Single slope ADC with reset counting for FeFET-based in-memory computing. In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5 (IEEE, 2024).
42. Zhao, C., Fang, J., Jiang, J., Xue, X. & Zeng, X. Light-CIM: a light-weight ADC/DAC-fewer RRAM CIM DNN accelerator with fully-analog tiles and non-ideality-aware algorithm for consumer electronics. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **44**, 602–612 (2024).
43. Nasiri, H., Li, C. & Zhang, L. Machine-learning-based SAR ADC featuring smart range detection for portable voice-activated IoT devices. In *IEEE Internet Things J.*, **1** (IEEE, 2025).
44. Oh, S. et al. Energy-efficient Mott activation neuron for full-hardware implementation of neural networks. *Nat. Nanotechnol.* **16**, 680–687 (2021).
45. Li, Y. et al. Low-voltage ultrafast nonvolatile memory via direct charge injection through a threshold resistive-switching layer. *Nat. Commun.* **13**, 4591 (2022).
46. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R. & Eleftheriou, E. Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **15**, 529–544 (2020).
47. Ning, H. et al. An in-memory computing architecture based on a duplex two-dimensional material structure for in situ machine learning. *Nat. Nanotechnol.* **18**, 493–500 (2023).
48. Yao, B. W. et al. Non-volatile electrolyte-gated transistors based on graphdiyne/MoS₂ with robust stability for low-power neuromorphic computing and logic-in-memory. *Adv. Funct. Mater.* **31**, 2100069 (2021).
49. Zhang, Z. C. et al. An ultrafast nonvolatile memory with low operation voltage for high-speed and low-power applications. *Adv. Funct. Mater.* **31**, 2102571 (2021).
50. Pi, L. et al. Broadband convolutional processing using band-alignment-tunable heterostructures. *Nat. Electron.* **5**, 248–254 (2022).
51. Zhou, F. et al. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat. Nanotechnol.* **14**, 776–782 (2019).
52. Wan, T. et al. In-sensor computing: materials, devices, and integration technologies. *Adv. Mater.* **35**, 2203830 (2023).
53. Chen, P. et al. Open-loop analog programmable electrochemical memory array. *Nat. Commun.* **14**, 6184 (2023).
54. Migliato Marega, G. et al. A large-scale integrated vector-matrix multiplication processor based on monolayer molybdenum disulfide memories. *Nat. Electron.* **6**, 991–998 (2023).
55. Wan, W. et al. A compute-in-memory chip based on resistive random-access memory. *Nature* **608**, 504–512 (2022).
56. Joshi, V. et al. Accurate deep neural network inference using computational phase-change memory. *Nat. Commun.* **11**, 2473 (2020).
57. Su, X., Guo, Z., Zhang, G., Tsai, H. & Li, N. Progress and challenges of phase change memory for in-memory computing. *Curr. Opin. Solid State Mater. Sci.* **37**, 101225 (2025).
58. Bonnet, D. et al. Bringing uncertainty quantification to the extreme-edge with memristor-based Bayesian neural networks. *Nat. Commun.* **14**, 7530 (2023).
59. Nandakumar, S. et al. A phase-change memory model for neuromorphic computing. *J. Appl. Phys.* **124**, 152135 (2018).
60. Liu, L. et al. Uniform nucleation and epitaxy of bilayer molybdenum disulfide on sapphire. *Nature* **605**, 69–75 (2022).
61. Wang, Y. et al. Ultraflat single-crystal hexagonal boron nitride for wafer-scale integration of a 2D-compatible high- κ metal gate. *Nat. Mater.* **23**, 1495–1501 (2024).
62. Wang, L. et al. Epitaxial growth of a 100-square-centimetre single-crystal hexagonal boron nitride monolayer on copper. *Nature* **570**, 91–95 (2019).
63. Neilson, K. M. et al. Toward mass production of transition metal dichalcogenide solar cells: scalable growth of photovoltaic-grade multilayer WSe₂ by tungsten selenization. *ACS Nano* **18**, 24819–24828 (2024).
64. Liu, L. et al. A mass transfer technology for high-density two-dimensional device integration. *Nat. Electron.* **8**, 135–146 (2025).
65. Lu, D. et al. Monolithic three-dimensional tier-by-tier integration via van der Waals lamination. *Nature* **630**, 340–345 (2024).
66. Kwon, J. et al. 200-mm-wafer-scale integration of polycrystalline molybdenum disulfide transistors. *Nat. Electron.* **7**, 356–364 (2024).
67. Peng, R. et al. Programmable graded doping for reconfigurable molybdenum ditelluride devices. *Nat. Electron.* **6**, 852–861 (2023).
68. Tsai, M.-Y. et al. A reconfigurable transistor and memory based on a two-dimensional heterostructure and photoinduced trapping. *Nat. Electron.* **6**, 755–764 (2023).
69. Wang, Z. et al. Reconfigurable quasi-nonvolatile memory/sub-thermionic FET functions in ferroelectric-2D semiconductor vdW architectures. *Adv. Mater.* **34**, 2200032 (2022).
70. Wang, Y. et al. Reconfigurable sensing-memory-processing and logical integration within 2D ferroelectric optoelectronic transistor for CMOS-compatible bionic vision. *Adv. Funct. Mater.* **34**, 2400039 (2024).
71. Xue, F. et al. Giant ferroelectric resistance switching controlled by a modulatory terminal for low-power neuromorphic in-memory computing. *Adv. Mater.* **33**, 2008709 (2021).
72. Kumar, M., Kim, U., Lee, W. & Seo, H. Ultrahigh-speed in-memory electronics enabled by proximity-oxidation-evolved metal oxide redox transistors. *Adv. Mater.* **34**, 2200122 (2022).

73. Yang, Q. et al. Controlled optoelectronic response in van der Waals heterostructures for in-sensor computing. *Adv. Funct. Mater.* **32**, 202207290 (2022).
74. Lee, S., Peng, R., Wu, C. & Li, M. Programmable black phosphorus image sensor for broadband optoelectronic edge computing. *Nat. Commun.* **13**, 1485 (2022).
75. Li, T. et al. Reconfigurable, non-volatile neuromorphic photovoltaics. *Nat. Nanotechnol.* **18**, 1303–1310 (2023).
76. Li, G. et al. Interface-engineered non-volatile visible-blind photo-detector for in-sensor computing. *Nat. Commun.* **16**, 57 (2025).
77. Lin, H. et al. In situ training of an in-sensor artificial neural network based on ferroelectric photosensors. *Nat. Commun.* **16**, 421 (2025).
78. Gong, Y. et al. Reconfigurable and nonvolatile ferroelectric bulk photovoltaics based on 3R-WS₂ for machine vision. *Nat. Commun.* **16**, 230 (2025).
79. Wang, Y. et al. A three-dimensional neuromorphic photosensor array for nonvolatile in-sensor computing. *Nano Lett.* **23**, 4524–4532 (2023).
80. Huang, H. et al. Fully integrated multi-mode optoelectronic memristor array for diversified in-sensor computing. *Nat. Nanotechnol.* **20**, 93–103 (2025).
81. Huang, H. et al. In-sensor compressing via programmable optoelectronic sensors based on van der Waals heterostructures for intelligent machine vision. *Nat. Commun.* **16**, 3836 (2025).
82. Li, Z. et al. Crossmodal sensory neurons based on high-performance flexible memristors for human-machine in-sensor computing system. *Nat. Commun.* **15**, 7275 (2024).
83. Li, Q. et al. High-performance ferroelectric field-effect transistors with ultra-thin indium tin oxide channels for flexible and transparent electronics. *Nat. Commun.* **15**, 2686 (2024).
84. Wang, H. et al. Silicon-compatible ferroelectric tunnel junctions with a SiO₂/Hf_{0.5}Zr_{0.5}O₂ composite barrier as low-voltage and ultra-high-speed memristors. *Adv. Mater.* **36**, 2211305 (2024).
85. Park, S.-O., Jeong, H., Park, J., Bae, J. & Choi, S. Experimental demonstration of highly reliable dynamic memristor for artificial neuron and neuromorphic computing. *Nat. Commun.* **13**, 2888 (2022).
86. Surekcigil Pesch, I., Bestelink, E., de Sagazan, O., Mehonic, A. & Sporea, R. A. Multimodal transistors as ReLU activation functions in physical neural network classifiers. *Sci. Rep.* **12**, 670 (2022).
87. Li, K. et al. Configurable activation function realized by non-linear memristor for neural network. *AIP Adv.* **10**, 085207 (2020).
88. Oh, J. et al. Preventing vanishing gradient problem of hardware neuromorphic system by implementing imidazole-based memristive ReLU activation neuron. *Adv. Mater.* **35**, 2300023 (2023).
89. Yang, Y. et al. Firing feature-driven neural circuits with scalable memristive neurons for robotic obstacle avoidance. *Nat. Commun.* **15**, 4318 (2024).
90. Xiao, Y. et al. Bio-plausible reconfigurable spiking neuron for neuromorphic computing. *Sci. Adv.* **11**, eadr6733 (2025).
91. Li, D. et al. Double-opponent spiking neuron array with orientation selectivity for encoding and spatial-chromatic processing. *Sci. Adv.* **11**, eadt3584 (2025).
92. Wang, Y. et al. A biologically inspired artificial neuron with intrinsic plasticity based on monolayer molybdenum disulfide. *Nat. Electron.* **8**, 680–688 (2025).
93. Li, J. et al. An ion-electronic hybrid artificial neuron with a widely tunable frequency. *Nat. Commun.* **16**, 7911 (2025).
94. Li, J. et al. Hardware-driven nonlinear activation for stochastic computing based deep convolutional neural networks. In *2017 International Joint Conference on Neural Networks (IJCNN)*, 1230–1236 (IEEE, 2017).
95. Leboeuf, K., Namin, A. H., Muscedere, R., Wu, H. & Ahmadi, M. High speed VLSI implementation of the hyperbolic tangent sigmoid function. In *2008 Third International Conference Convergence and Hybrid Information Technology*, 1, 1070–1073 (IEEE Computer Society, 2008).
96. Sun, H. et al. A universal method of linear approximation with controllable error for the efficient implementation of transcendental functions. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**, 177–188 (2019).
97. Qin, Z. et al. A novel approximation methodology and its efficient VLSI implementation for the sigmoid function. *IEEE Trans. Circuits Syst. II Express Briefs* **67**, 3422–3426 (2020).
98. Kim, D. & Lee, B. -g Energy-and area-efficient CMOS neuron and max pooling circuit for RRAM-based CNN accelerators. *IEEE Access* **13**, 84329–84340 (2025).
99. Pan, Z., Gu, Z., Jiang, X., Zhu, G. & Ma, D. A modular approximation methodology for efficient fixed-point hardware implementation of the sigmoid function. *IEEE Trans. Ind. Electron.* **69**, 10694–10703 (2022).
100. Hajduk, Z. & Dec, G. R. Very high accuracy hyperbolic tangent function implementation in FPGAs. *IEEE Access* **11**, 23701–23713 (2023).
101. Roshan, A. D., Guha, P. & Trivedi, G. Hardware-optimized regression tree-based Sigmoid and Tanh functions for machine learning applications. *IEEE Trans. Circuits Syst. II Express Briefs* **72**, 283–287 (2024).
102. Amin, M. H., Elbity, M., Mohammadi, M. & Zand, R. MRAM-based analog sigmoid function for in-memory computing. In *Proc. Great Lakes Symposium on VLSI 2022*, 319–323 (Association for Computing Machinery, 2022).
103. Gross, A. L. et al. Self-heating electrochemical memory for high-precision analog computing. *arXiv preprint arXiv:2505.15936* (2025).
104. Lee, C. et al. Improved on-chip training efficiency at elevated temperature and excellent inference accuracy with retention (> 10⁹ s) of Pr_{0.7}Ca_{0.3}MnO_{3-x} ECRAM synapse device for hardware neural network. In *2021 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2021).
105. Xiao, T. P. et al. The effect of heavy ion strikes on charge trap memory arrays with analog state programmability. *IEEE Trans. Nucl. Sci.* **72**, 1375–1383 (2025).
106. Milo, V. et al. Optimized programming algorithms for multilevel RRAM in hardware neural networks. In *2021 IEEE International Reliability Physics Symposium (IRPS)*, 1–6 (IEEE, 2021).

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2021YFA1400601 (J.T.)), the National Natural Science Foundation of China (62274119 (X.C.)), Basic Research Development Program of Suzhou (SJC2023004 (L.K.)), and the Shenzhen Science and Technology Program (JCYJ20240813160206009 (Z.C.)), and the Fundamental Research Funds for the Central Universities (010-63253118 (X.C.)). The authors are grateful for the technical support for Nano-X from Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences (SINANANO).

Author contributions

Z.Z., L.K., and X.C. designed the project, and X.C., J.T., and T.L. supervised this project. Z.Z., Y.L., and J.Y. fabricated the device and performed the electrical/optoelectronic measurements. F.W. prepared the GDYO film. S.S. and H.Q. performed the characterization. X.C. wrote the manuscript, Z.L. and Z.C. revised the manuscript. All the authors discussed the results and commented on 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/s41467-026-68402-7>.

Correspondence and requests for materials should be addressed to Lixing Kang, Zhi-Bo Liu, Jian-Guo Tian or Xu-Dong Chen.

Peer review information *Nature Communications* thanks Suhas Kumar 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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2026