Advanced Materials Review paper: Neuromorphic Engineering: From Biological to Spike-Based Hardware Nervous Systems

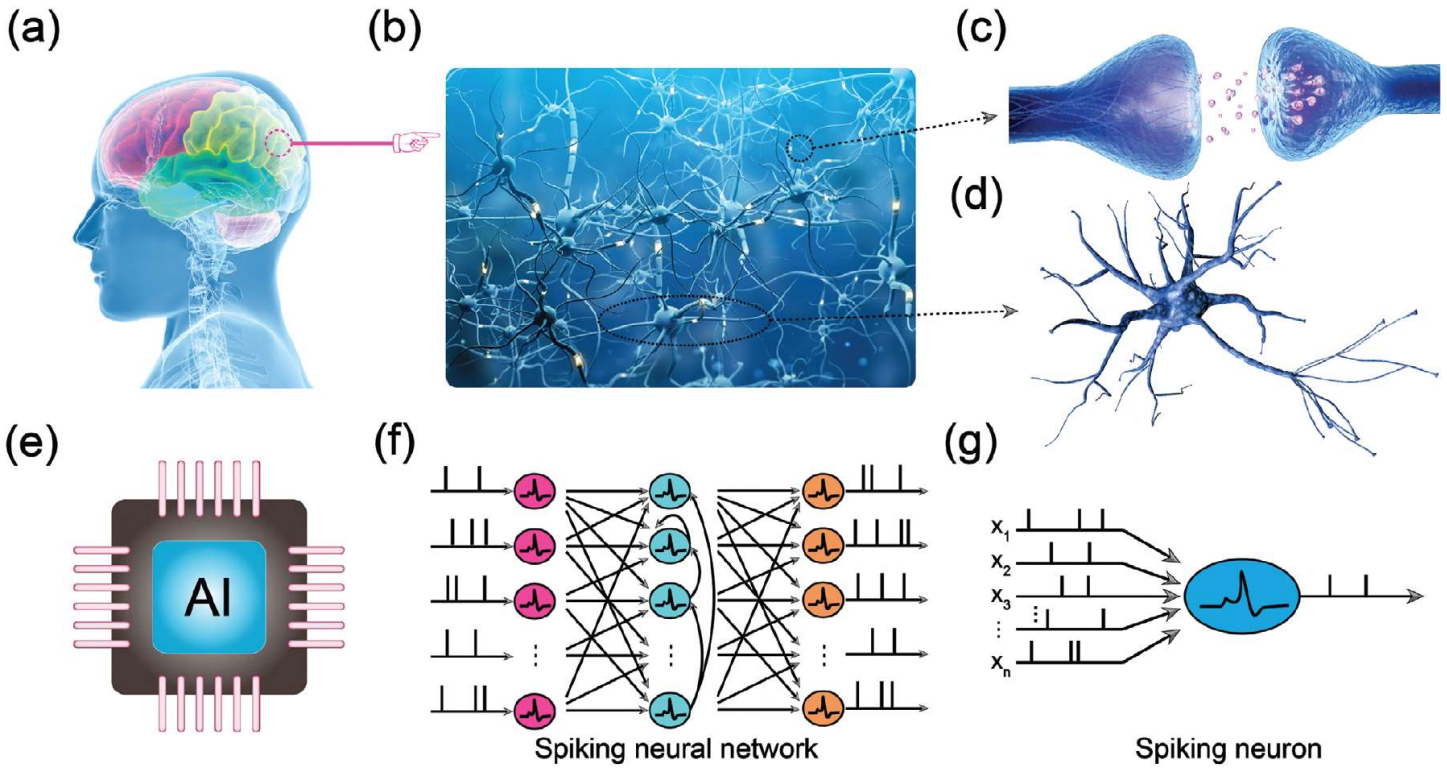


Figure 1. Schematic diagram of biological and artificial computing systems. a) The human brain. b) The biological neural network. c) A biological synapse. d) A biological neuron. e) An AI chip. f) Spiking neural networks. g) An artificial spiking neuron. a) Adapted with permission.[231] Copyright 2020, PAIXIN. b) Adapted with permission.[232] Copyright 2020, PAIXIN.

Human brain has massive parallelism, extremely low power consumption, superior fault tolerance, strong robustness

ANN:
Gen1-> MLP
Gen2-> activation functions (SVM, CNN, RNN)
Gen3-> spiking neurons

How the biological neurons work

From synapse

To other neurons' dendrites through synaptic transmission

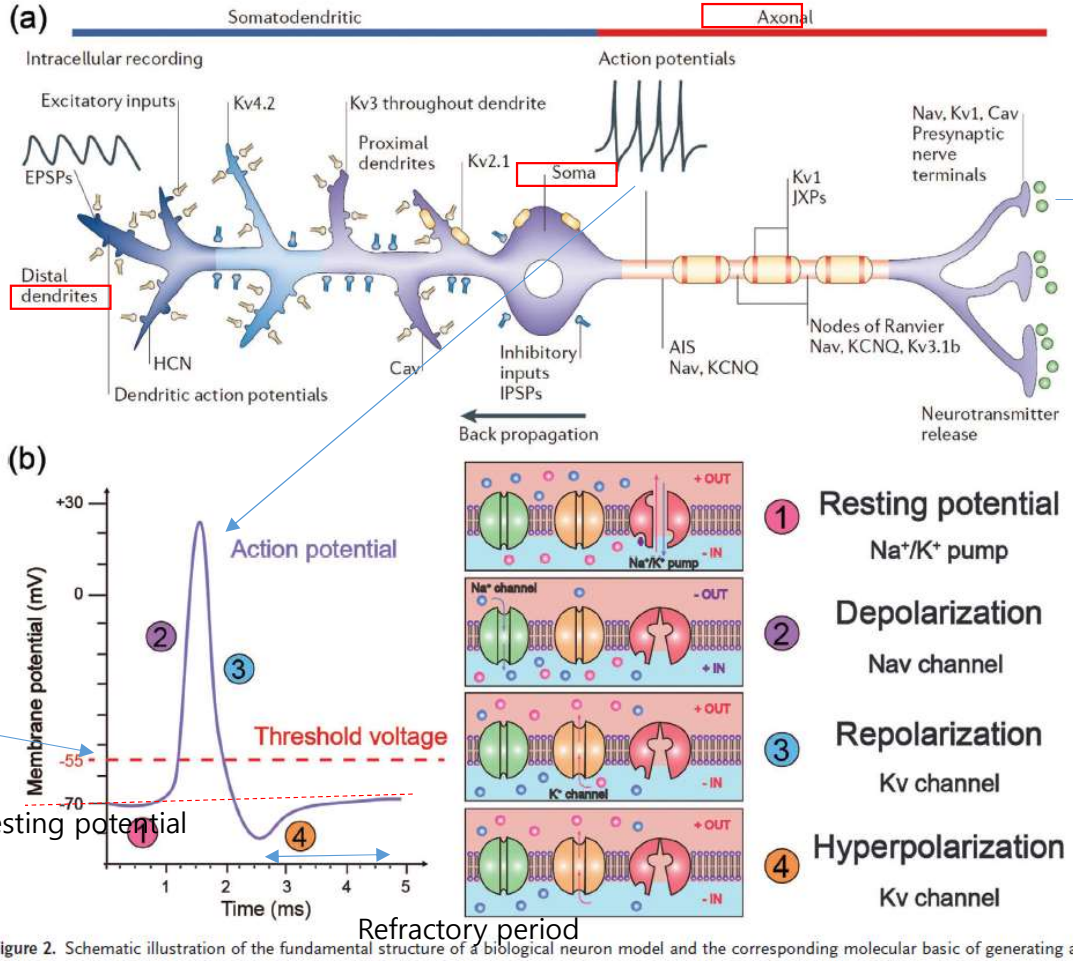


Figure 2. Schematic illustration of the fundamental structure of a biological neuron model and the corresponding molecular basic of generating an action potential. a) A neuron receives stimulations at the dendrites and generates action potentials propagating along the axon when the membrane potential overshoots a certain threshold voltage. In this process, the ion channels in different parts of the neuron interact to participate in the generation of action potentials. b) The generation of an action potential which can be divided into four stages, including resting potential, depolarization, repolarization and hyperpolarization. The Na^+/K^+ pump is responsible for restoring resting potential. The activation of Nav channel leads to depolarization. The Kv channel is responsible for repolarization and hyperpolarization. a) Adapted with permission.[36] Copyright 2006, Springer Nature.

Crossing the threshold creates action potential

Membrane potential goes up as stimuli are summed up.

Neuron's main structure:
Dendrite, soma, axon

Neuron connects with other neurons through synapses.

Synaptic Transmission

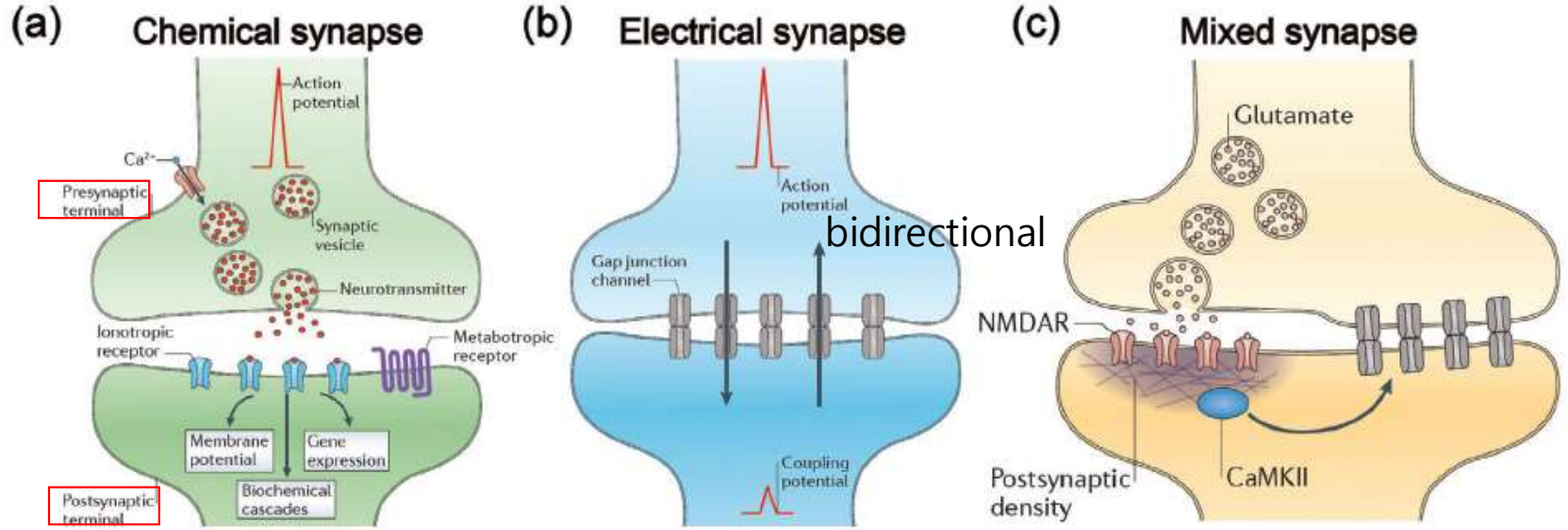
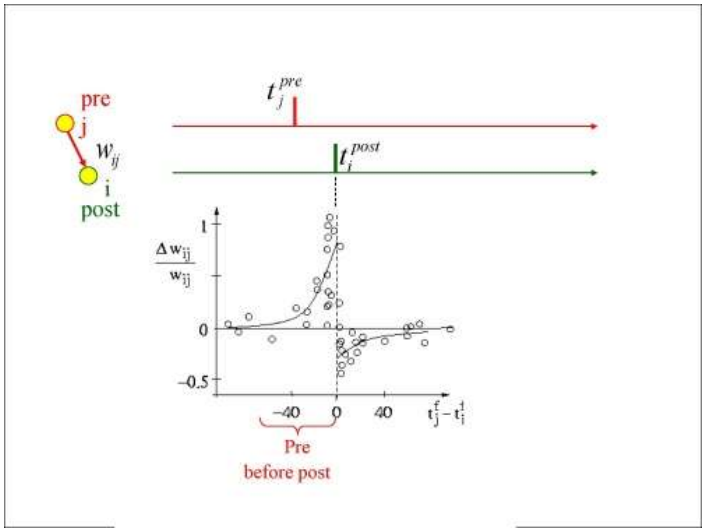


Figure 3. Main modalities of synaptic transmissions. a) chemical synapse. The arrival of action potential results in the activation of voltage-gated Ca^+ channels, promoting the probabilistic release of neurotransmitters by exocytosis from presynaptic membrane. The ionotropic and metabotropic receptors on the postsynaptic membrane can detect and translate the information carried by neurotransmitters into different postsynaptic behaviors, varying from changes in membrane potential to gene expression. b) Electrical synapse. Electrical transmission is conducted by gap junctions (some clusters of intercellular channels) between two adjacent cells. The transmission is bidirectional: when an action potential is transmitted from pre-synapse to postsynapse, the postsynaptic resting potential propagates concurrently to the pre-synapse. c) Mixed synapse. Chemical and electrical transmission coexist at mixed synapses. Chemical synapses (such as glutamate-based) influence the connective strength of electrical synapses by activating the NMDA receptors and CaMKII. a–c) Adapted with permission.[55] Copyright 2014, Springer Nature.

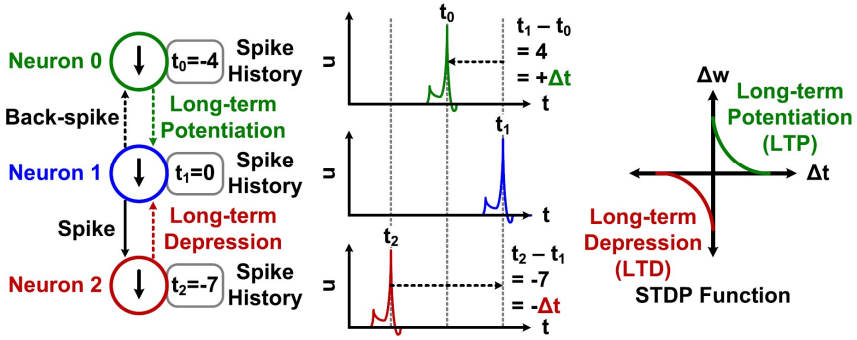
Spike timing dependent plasticity (STDP)



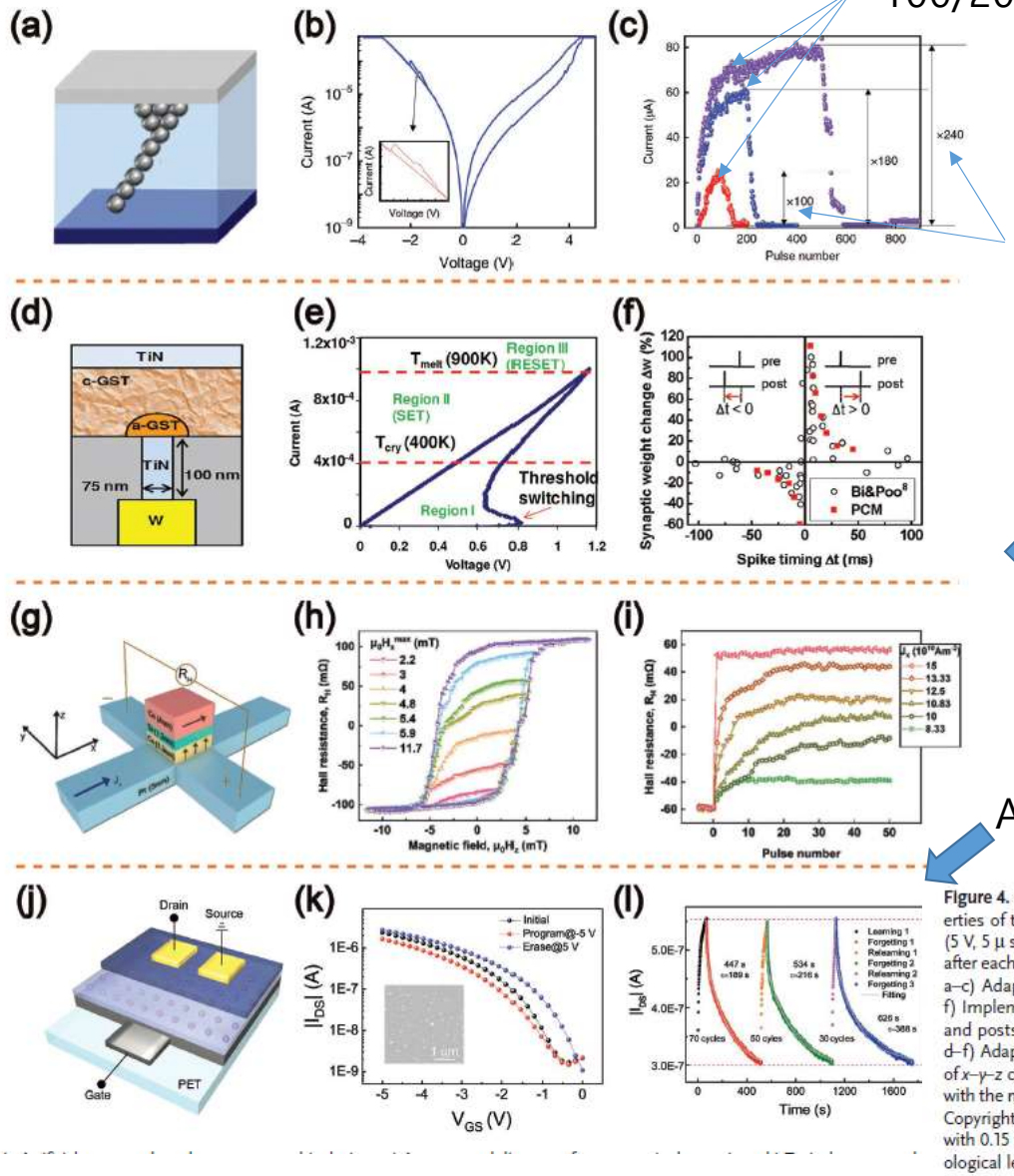
$$\Delta w_j = \sum_{f=1}^{N} \sum_{n=1}^{N} W(t_i^n - t_j^f)$$

$$W(x) = A_+ \exp(-x/\tau_+) \qquad for \quad x > 0$$

$$W(x) = -A_- \exp(x/\tau_-) \qquad for \quad x < 0$$



Artificial Synapses



100/200/500 set followed by the same number reset

Memrister characteristics and potentiation with spikes

On, off ratio

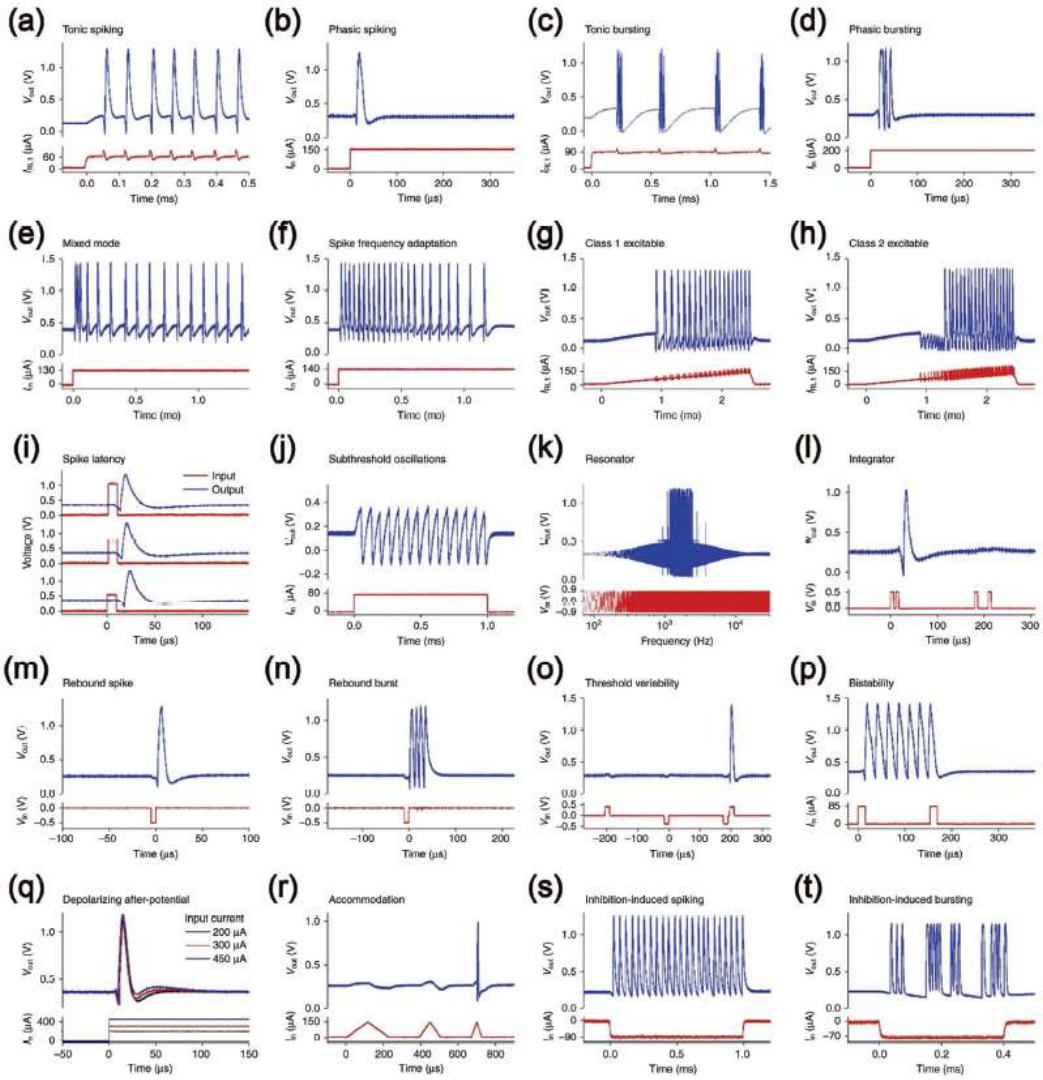
Phase change memory used to implement STDP

A type of MRAM

A synaptic transistor

Figure 4. Artificial synapses based on neuromorphic devices. a) A conceptual diagram of a two-terminal memristor. b) Typical current–voltage (*I*–*V*) properties of the memristor. c) Conductance potentiation and depression of the synaptic device. A pulse train comprises sequential 100/200/500 set pulses (5 V, 5 μs) followed by sequential 100/200/500 reset pulses (−3 V, 5 μs) was applied to the device, and the response current was read at (2 V, 1 ms) pulse after each set/reset pulse. The analog on/off ratios are × 100, × 180 and × 240 at 100-100, 200-200, and 500-500 potentiation-depression pulses, respectively. a–c) Adapted with permission.[14] Copyright 2018, Springer Nature. d) Schematic diagram of a phase change cell. e) *I*–*V* characteristics of a PCM cell. f) Implementation of STDP with PCM cells. The change of synaptic weights is described as a function of the relative time difference between presynaptic and postsynaptic spikes. Experimental data is shown together with the one measured by Bi and Poo in biological hippocampal glutamatergic synapses. d–f) Adapted with permission.[15] Copyright 2012, American Chemical Society. g) Device structure of a double magnetic layered system and the definition of *x*–*y*–*z* coordinates for experimental measurement. h) R_H–H_z loops with $\mu_0 H_z$ sweeping from −11.7 mT to $\mu_0 H_z^{max}$ and then back to −11.7 mT. i) R_H varies with the number of current pulses (total 50) in different constant magnitudes J_x. The R_H is measured 2 s after each pulse. g–i) Adapted with permission.[123] Copyright 2019, Wiley-VCH. j) Schematic of a bottom-gate top-contact flexible synaptic transistor. k) Typical transfer plot of the flexible synaptic transistor with 0.15 wt% C_{60}. Inset: Corresponding top-view SEM image of the C_{60}/PMMA hybrid film utilized in the synaptic transistor. l) Implementation of physiological learning, forgetting, and relearning processes with the proposed C_{60}-based device. j–l) Adapted with permission.[17] Copyright 2018, Wiley-VCH.

Biological Neuron Models

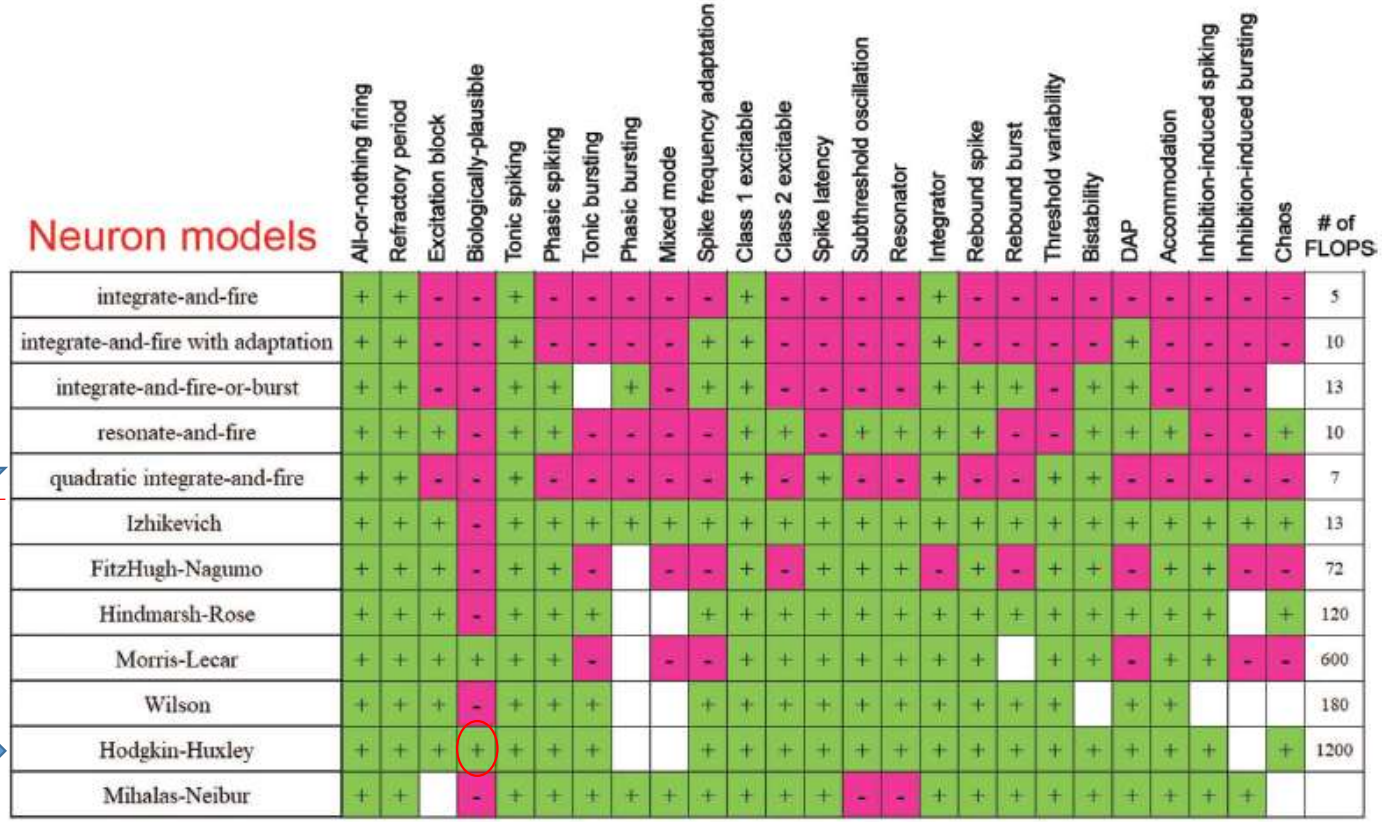


Different neurons exhibit different behaviors

Figure 5. Schematic illustration of the 20 common biological spiking behaviors experimentally demonstrated in memristor based neurons. a) Tonic spiking. b) Phasic spiking. c) Tonic bursting. d) Phasic bursting. e) Mixed mode. f) Spike frequency adaptation. g) Class 1 excitable. h) Class 2 excitable. i) Spike latency. j) Subthreshold oscillations. k) Resonator. l) Integrator. m) Rebound spike. n) Rebound burst. o) Threshold variability. p) Bistability. q) Depolarizing after-potential. r) Accommodation. s) Inhibition-induced spiking. t) Inhibition-induced bursting. a–s) Adapted under the terms of the CC-BY Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0).[24] Copyright 2018, The Authors, published by Springer Nature.

Good for digital Hardware (Leaky integrate and fire, (LIF))

Well known model, but too complex, needs analog processing



Neuron models	All-or-nothing firing	Refractory period	Excitation block	Biologically-plausible	Tonic spiking	Phasic spiking	Tonic bursting	Phasic bursting	Mixed mode	Spike frequency adaptation	Class 1 excitable	Class 2 excitable	Spike latency	Subthreshold oscillation	Resonator	Integrator	Rebound spike	Rebound burst	Threshold variability	Bistability	DAP	Accommodation	Inhibition-induced spiking	Inhibition-induced bursting	Chaos	# of FLOPS
integrate-and-fire	+	+	-	-	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	5
integrate-and-fire with adaptation	+	+	-	-	+	-	-	-	-	+	+	-	-	-	-	+	-	-	-	-	+	-	-	-	-	10
integrate-and-fire-or-burst	+	+	-	-	+	+		+	-	+	+	-	-	-	-	+	+	+	-	+	+	-	-	-		13
resonate-and-fire	+	+	+	-	+	+	-	-	-	-	+	+	-	+	+	+	+	-	-	+	+	+	-	-	+	10
quadratic integrate-and-fire	+	+	-	-	+	-	-	-	-	-	+	-	+	-	-	+	-	-	+	+	-	-	-	-	-	7
Izhikevich	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	13
FitzHugh-Nagumo	+	+	+	-	+	+	-		-	-	+	-	+	+	+	-	+	-	+	+	-	+	+	-	-	72
Hindmarsh-Rose	+	+	+	-	+	+	+			+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	120
Morris-Lecar	+	+	+	+	+	+	-		-	-	+	+	+	+	+	+	+		+	+	-	+	+	-	-	600
Wilson	+	+	+	-	+	+	+			+	+	+	+	+	+	+	+	+	+		+	+				180
Hodgkin-Huxley	+	+	+	+	+	+	+			+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	1200
Mihalas-Neibur	+	+		-	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+	+		

Figure 6. Comparison of spiking dynamics of different neuron models. # of FLOPS is the approximate number of floating-point operations required for the model simulation within 1 ms duration with a digital computer. (+), (–) and empty square represent the processed, missing, and unconfirmed properties of the corresponding neuron models, respectively. Adapted with permission.[141] Copyright 2004, IEEE.

Analog CMOS neurons

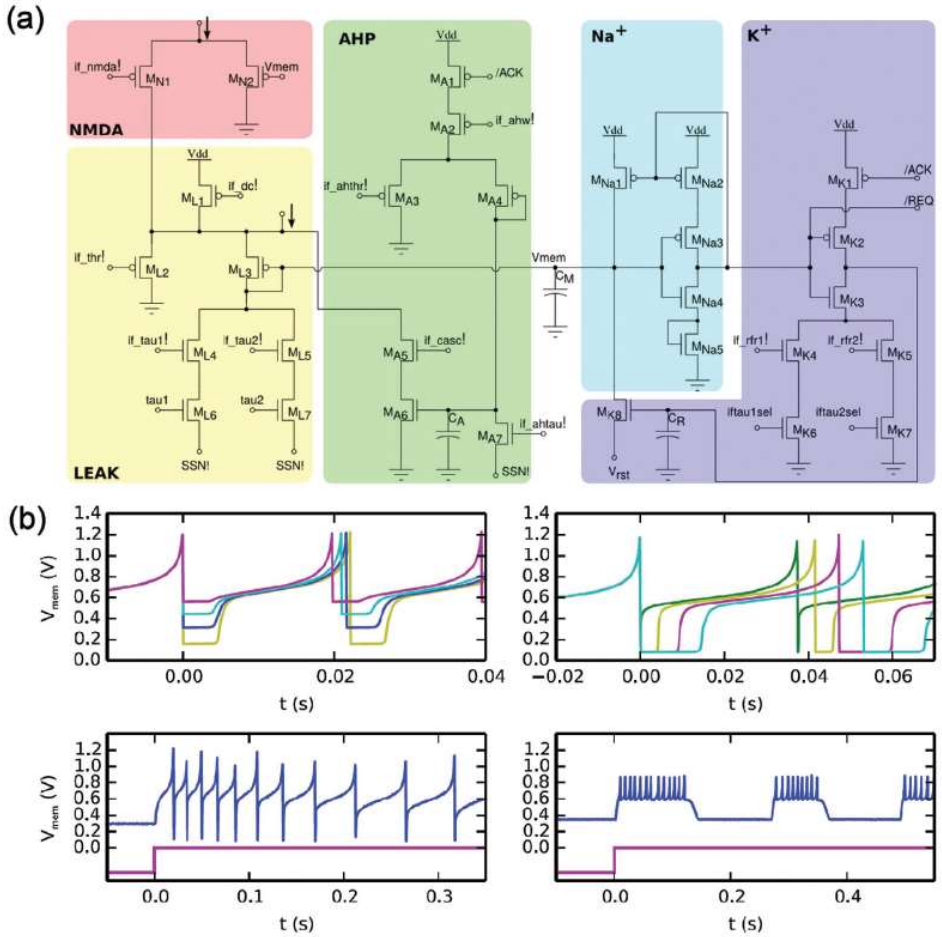


Figure 7. Implementation of analog CMOS-based neurons. a) The neuron model comprises 5 building blocks: a NMDA block, a LEAK block, an AHP block, a Na^+ channel block and a K^+ channel block. b) Configurable biologically-plausible spiking behaviors realized by the proposed neuron model, including tunable reset potential (upper-left), tunable refractory period duration (upper-right), spike-frequency adaptation behavior (lower-left) and bursting behavior (lower-right). a,b) Adapted under the terms of the CC-BY Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0).[169] Copyright 2015, The Authors, published by Frontiers.

Electrochemical metallization based Neurons (memrister neurons)

Current through this memrister acts as a spike if enough membrane potential is reached

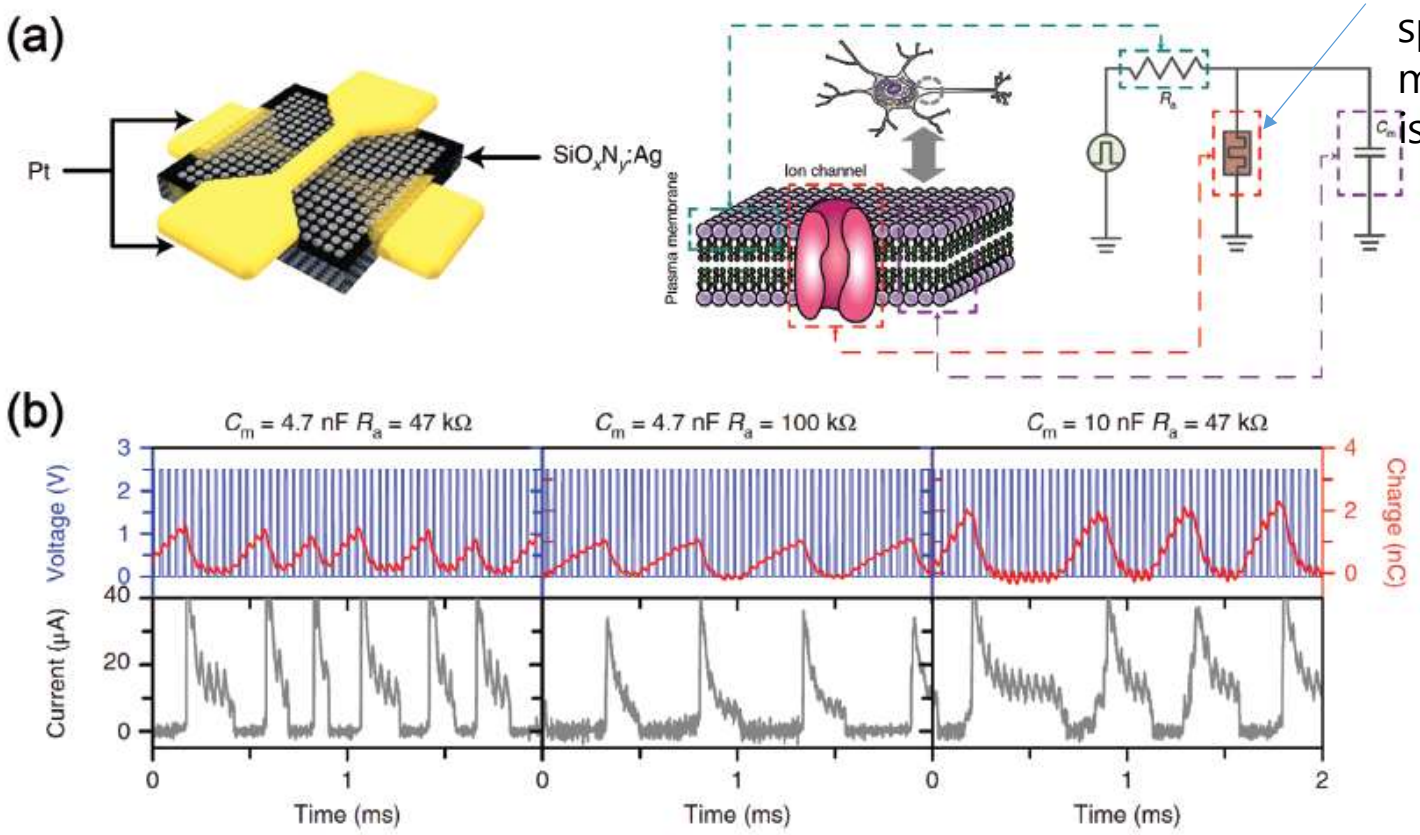


Figure 8. ECM-based neurons. a) Schematic illustration of the realization of an artificial LIF neuron. A cross-point diffusive memristor in a configuration of Pt/SiO_xN_y:Ag/Pt for artificial neuron (left). An ion channel on the cell membrane of a biological neuron (middle). The hardware implementation of LIF neuron (right), in which the memristor serves as the ion channel while the capacitor as cytomembrane. b) Controlling the response of the LIF neuron to multiple pulse trains by varying C_m and R_a. The current response across the memristor clearly depicts the firing of the neuron. a,b) Adapted with permission.[21] Copyright 2018, Springer Nature.

Valence change based Neurons (RRAM neurons)

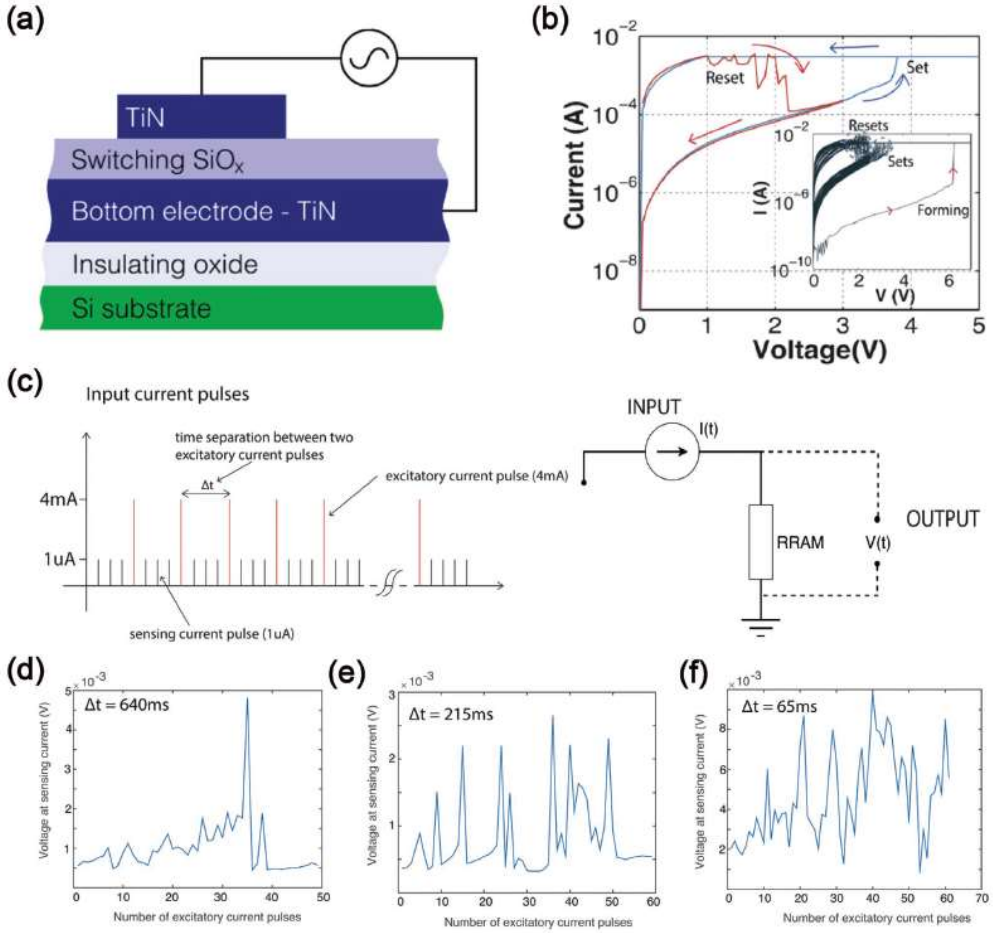


Figure 9. VCM-based neurons. a) Schematic diagram of a unipolar resistive random-access memory. b) Typical *I–V* curve of the device in (a). The right inset shows the resistive switching and forming process of the device. a,b) Adapted with permission.[177] Copyright 2015, AIP Publishing. c) Configuration of the input train, in which excitatory current spikes (4 mA) are separated by sensing current spikes (1 μA). d–f) Voltage response measured with the sensing current spikes after the excitatory one with time intervals of 640 ms (d), 215 ms (e), and 65 ms (f). The number of excitatory current pulses required to fire the neuron decreases with the shorter time interval. c–f) Adapted under the terms of the CC-BY Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0).[22] Copyright 2016, The Authors, published by Frontiers.

Valence change based Neurons (PCM based neurons)

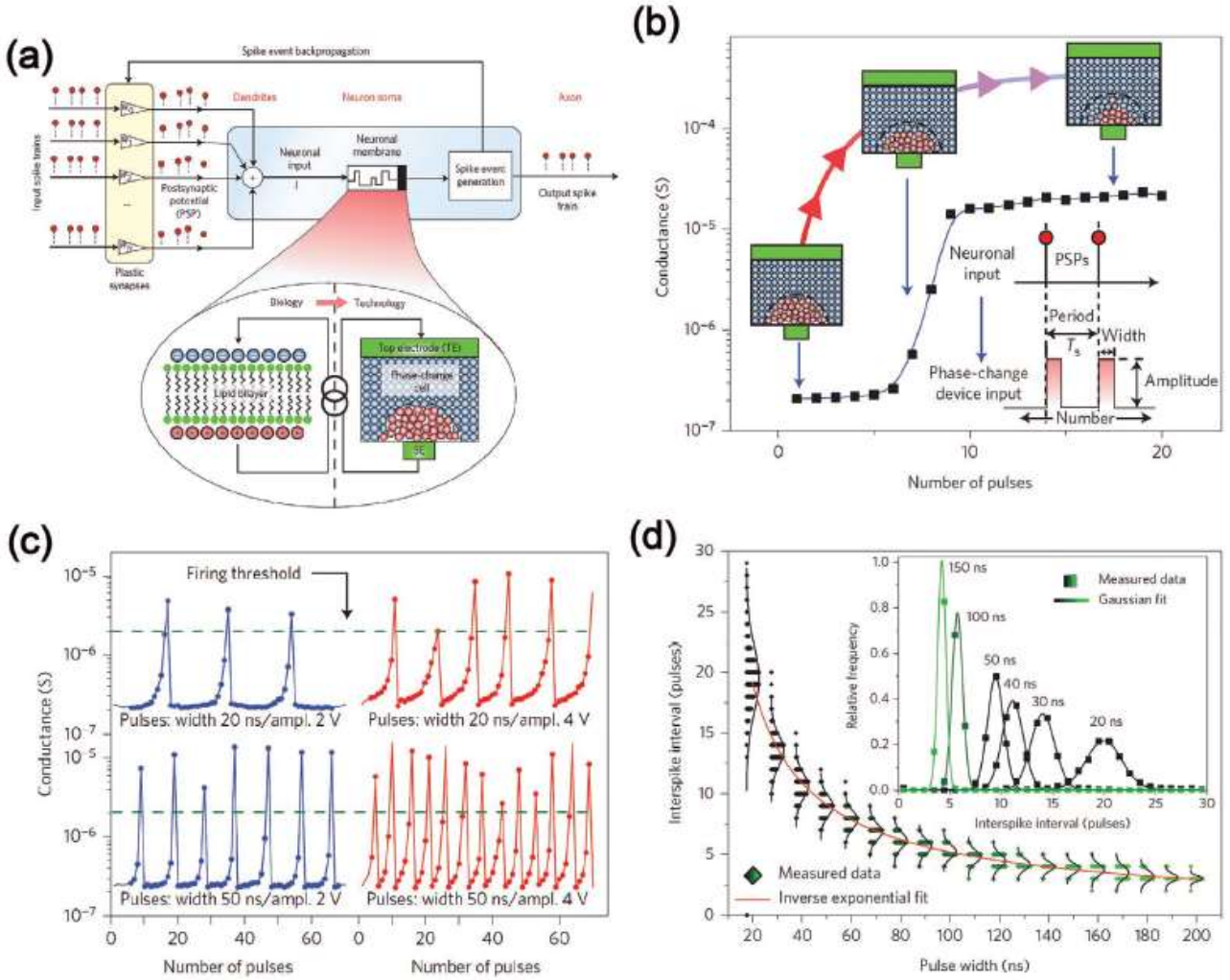


Figure 10. Phase-change-based neurons. a) Artificial neuron based on a phase-change memristor connects to a plastic synaptic input array. b) The conductance change of a mushroom-type phase-change memory in response to the crystallizing pulses. c) The integrate-and-fire dynamics of the proposed phase-change neuron. d) Stochastic firing behavior of the phase-change neuron. For a certain range of pulse widths, the interspike intervals are distributed normally at a fixed pulse amplitude of 2 V (1000 trails for each pulse width). a–d) Adapted with permission.[23] Copyright 2016, Springer Nature.

Insulatror to metal transition based neurons (Mott memory)

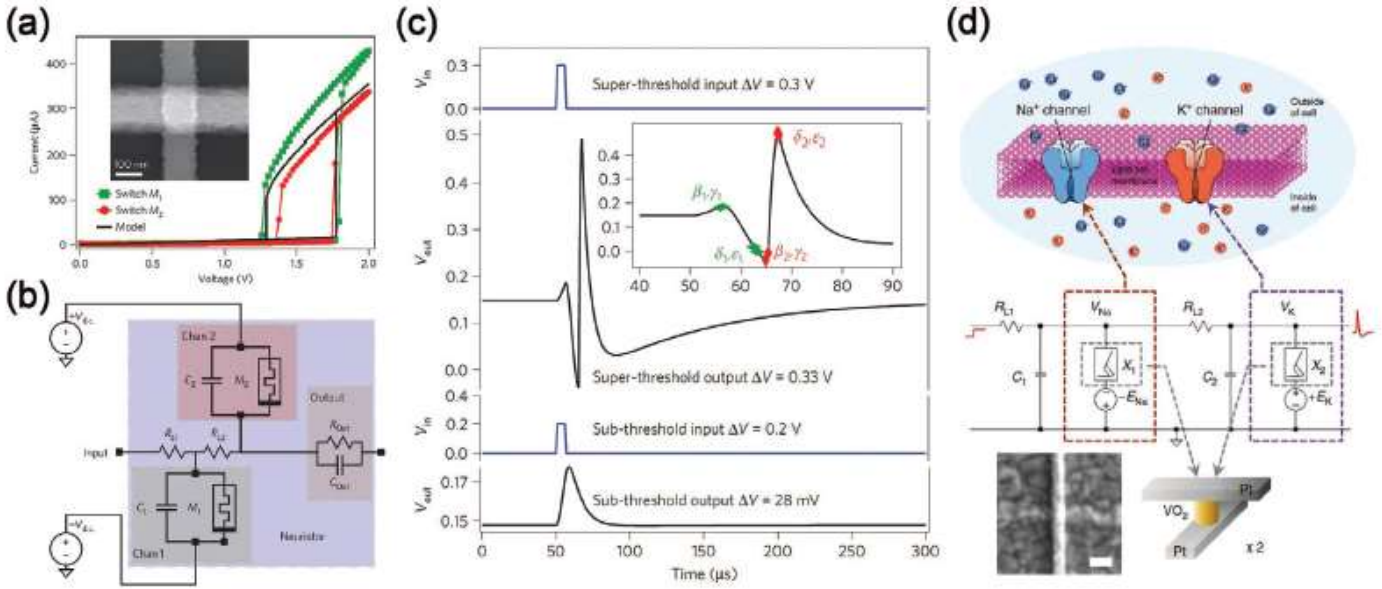


Figure 11. IMT-based neurons. a) Bistable *I–V* characteristics of two 110 × 110 nm^2 Mott memristors and that of model for simulation. b) Circuit diagram of the proposed neuristor. c) All-or-nothing spiking behavior of the neuristor. It demonstrates the output properties of the neuristor in response to a superthreshold pulse (0.3 V, 10 μs) and a subthreshold pulse (0.2 V, 10 μs). d) Implementation of a biologically plausible neuron with two Pt/VO_2/Pt Mott memristors. a–c) Adapted with permission.[185] Copyright 2012, Springer Nature. d) Adapted under the terms of the CC-BY Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0).[24] Copyright 2018, The Authors, published by Springer Nature.

Spin based neurons (MRAM based)

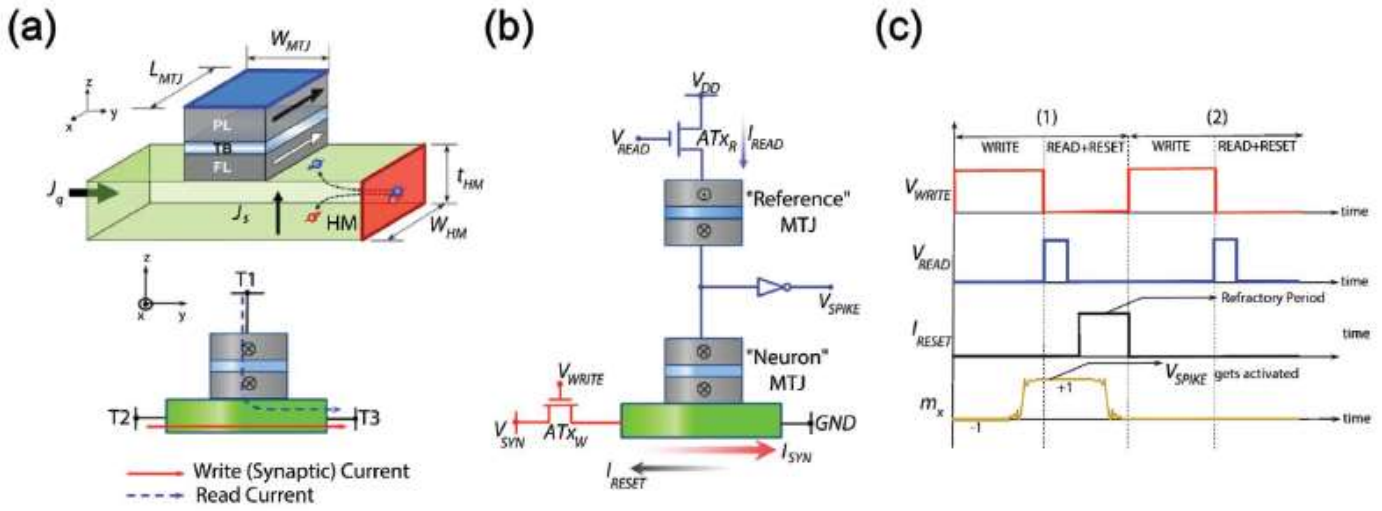


Figure 12. Spin-based neurons. a) Schematic illustration of a typical SOT-based device in which an MTJ is placed on a HM with high spin–orbit coupling. Input current flowing through the HM can effectively modulate the spin orientation of the FL, leading to the resistive change of the device. b) Circuit connection diagram of a stochastic neuron based on the spin-based devices. Read/write control transistors interface with the Neuron MTJ for the decouple of read/write current signals. When the neuron fires, a reset current IRESET is sent to reset the neuron MTJ to the P state. c) Two complete working cycles of the stochastic neuron in (b). a–c) Adapted with permission.[25] Copyright 2016, Springer Nature.

SNN hardware realization: input encoding

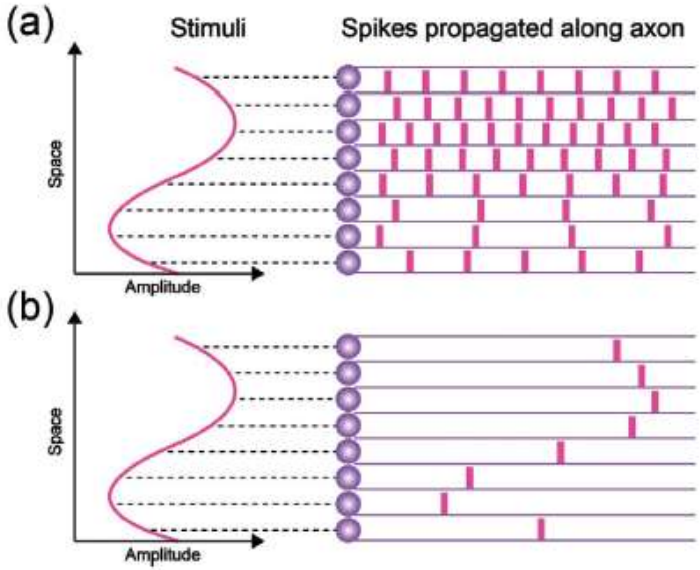


Figure 13. Data encoding schemes in SNNs. a) Rate-coding. A Higher input intensity is represented by higher spiking rate. b) Spatiotemporal-coding. Information is encoded into the precise relative spiking time between neurons. Neurons corresponding to higher intensity of stimuli spike earlier, while neurons connected to lower stimuli spike later.

IBM's all digital SNN using LIF: TrueNorth

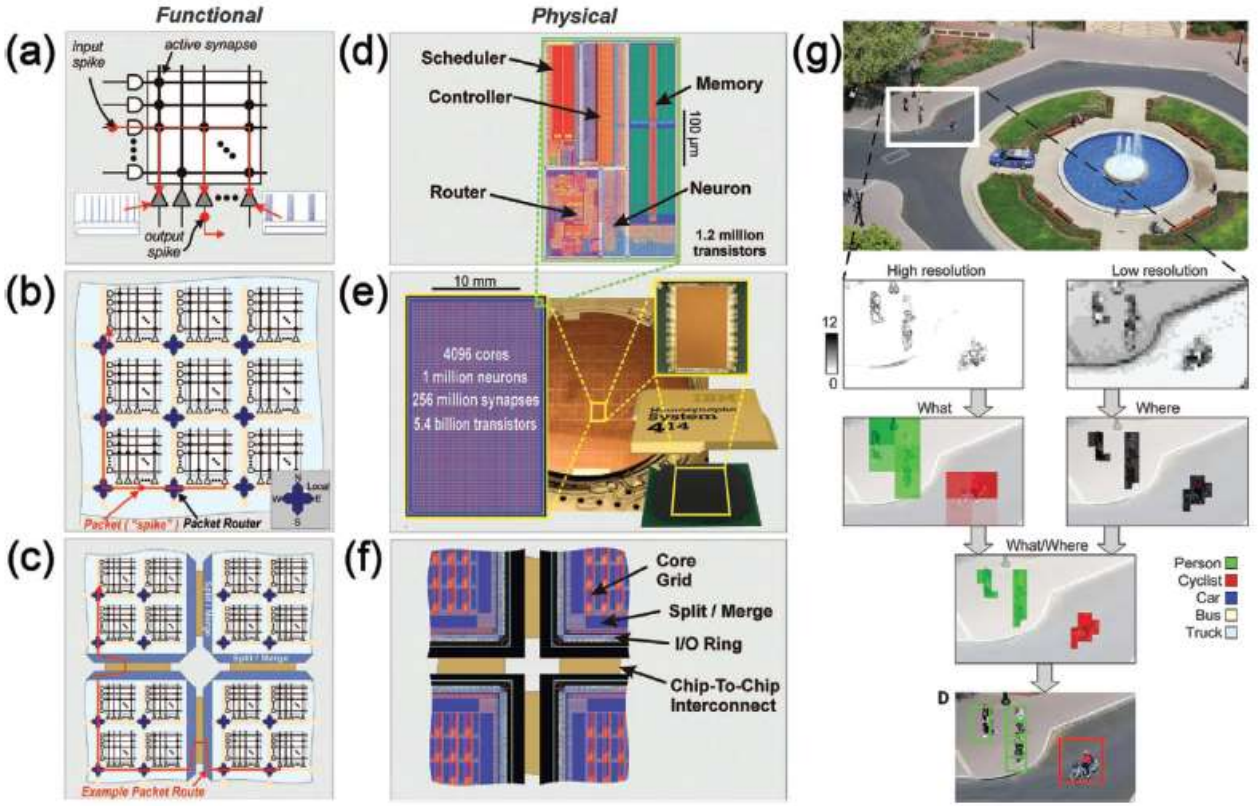


Figure 14. The TrueNorth neuromorphic computing system. a) Functional view of a neurosynaptic core which consists of axons (horizontal lines), programmable synapses (cross points), neuron inputs (vertical lines), and neurons (triangles). Neuronal dynamics can be programed individually. b) Functional chip architecture is a 2D core array, in which a mesh routing network is utilized to realize long-range connections. c) Routing network extends the scale of the network through peripheral merge and split blocks. d) Physical layout of the 28 nm CMOS based neurosynaptic core. e) Physical layout of the 64 × 64 core array, wafer and chip package. f) Peripheral circuits for multichip networks. g) Demonstration of real-time multiobject recognition based on TrueNorth platform. Video data set from a fixed camera is transduced into two spike-based parallel channels for labeling objects (high-solution channel) and locating salient objects (low-solution channel), respectively. The two pathways are associated with form a what–where map. After offline training, the chip can efficiently recognize different objects by reporting object bounding boxes. a–g) Adapted with permission.[211] Copyright 2014, The American Association for the Advancement of Science.

The chip contains 4096 neurosynaptic cores and 256 million synapses in total.

Intel Lohi (2018)/Lohi2(2021): SNN based

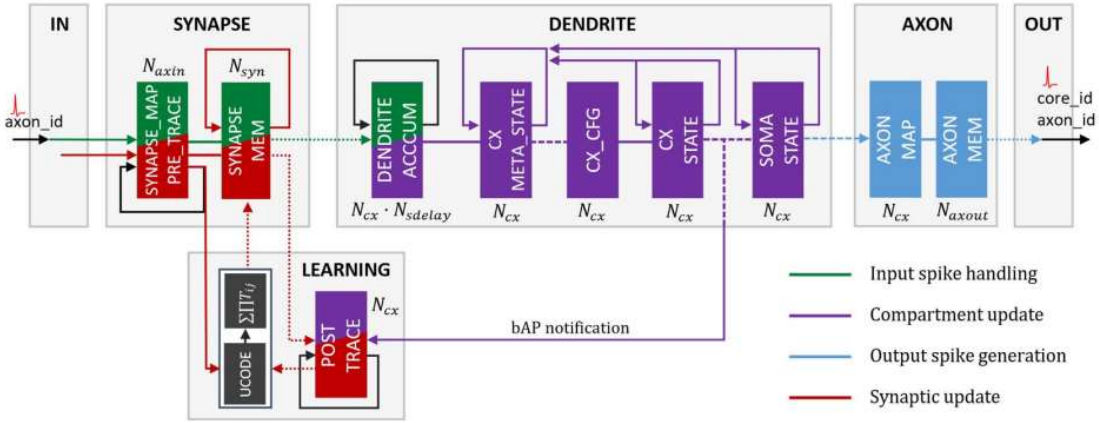
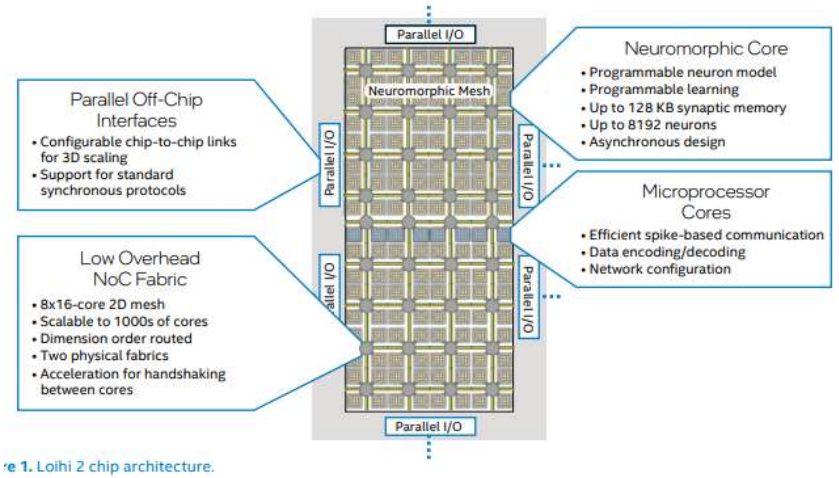
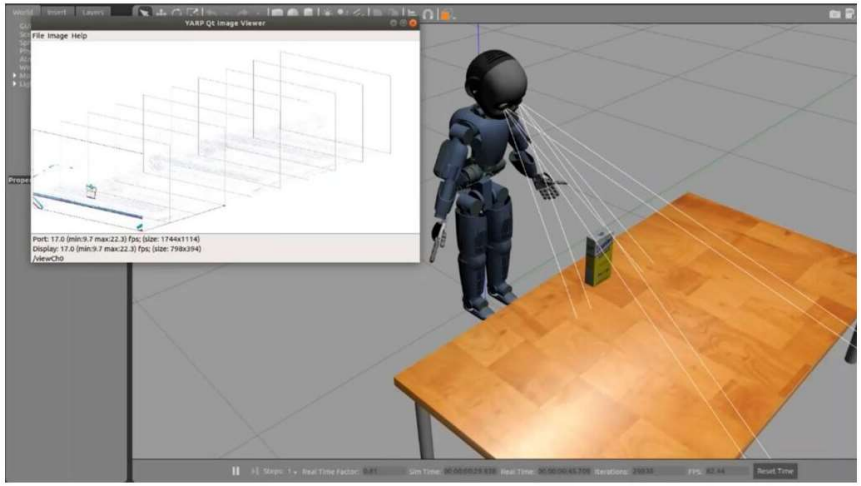


Figure 4. Core Top-Level Microarchitecture. The SYNAPSE unit processes all incoming spikes and reads out the associated synaptic weights from the memory. The DENDRITE unit updates the state variables *u* and *v* of all neurons in the core. The AXON unit generates spike messages for all fan-out cores of each firing neuron. The LEARNING unit updates synaptic weights using the programmed learning rules at epoch boundaries.



re 1. Loihi 2 chip architecture.



In a Simulated Setup, A Robot Actively Senses Objects By Moving Its Eyes (Event-Based Camera or Dynamic Vision Sensor) Generating "miscrosaccades" The Events Collected Are Used To Drive A Spiking Neural Network On The Lohi Chip. If The Object Or View Is New, Its SNN Representation Is Learned Or Updated. If The Object Is Known, It Is Recognized By The Network, And Respective Feedback Is Given To The User. (Credit: Intel Corporation)

Tianjic from Tsinghua University: ANN + SNN

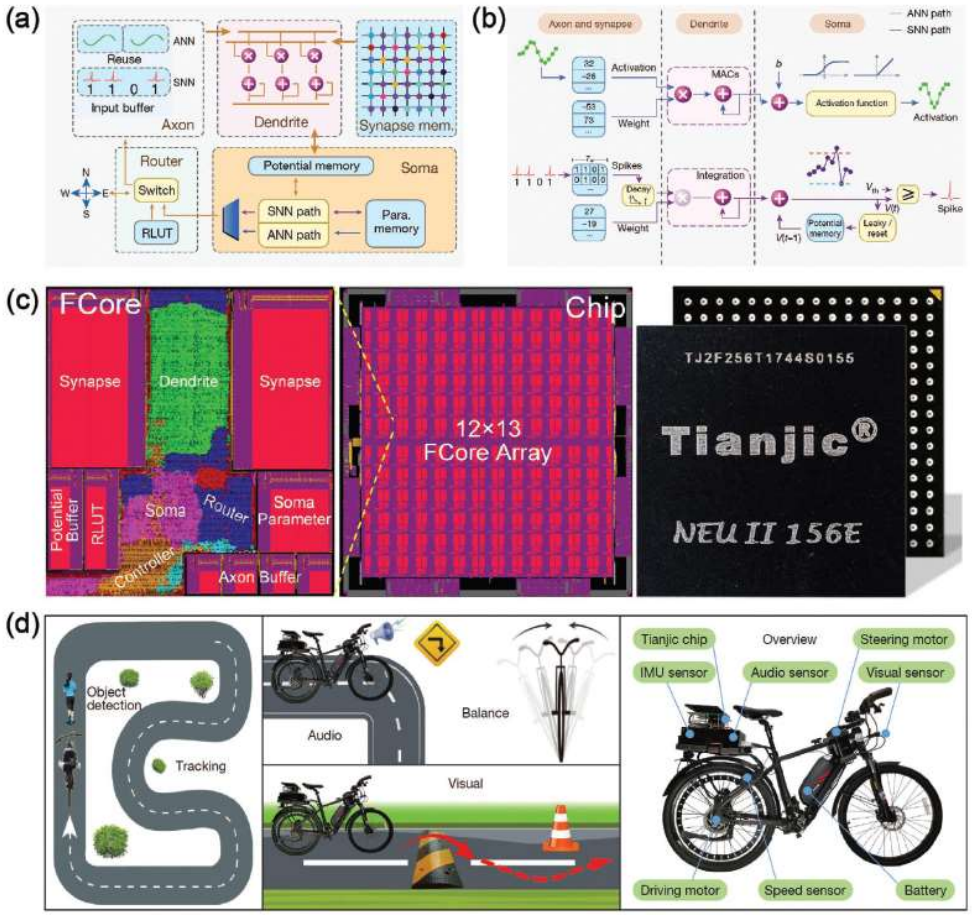


Figure 15. The Tianjic chip computing system. a) Schematic diagram of the configuration of a FCore. b) Comparison of information processing between the ANN neuron and the SNN neuron. $V(t)$ represents the membrane potential at time step t. V_{th} represents the firing threshold voltage. The numbers in blue boxes contain some examples of input activations/spikes and synaptic weights. c) Physical layout of the FCore and the Tianjic chip. d) Demonstration of unmanned bicycle driving based on a Tianjic chip. The tasks conducted in the experiment include object detection, tracking, audio recognition, balance control, obstacle avoidance, and so on. The bicycle system was equipped with multiple sensors, such as inertial measurement unit (IMU), gyroscope, audio sensor, visual sensor, motors, etc. a–d) Adapted with permission.[213] Copyright 2019, Springer Nature.

Tianjic : continued

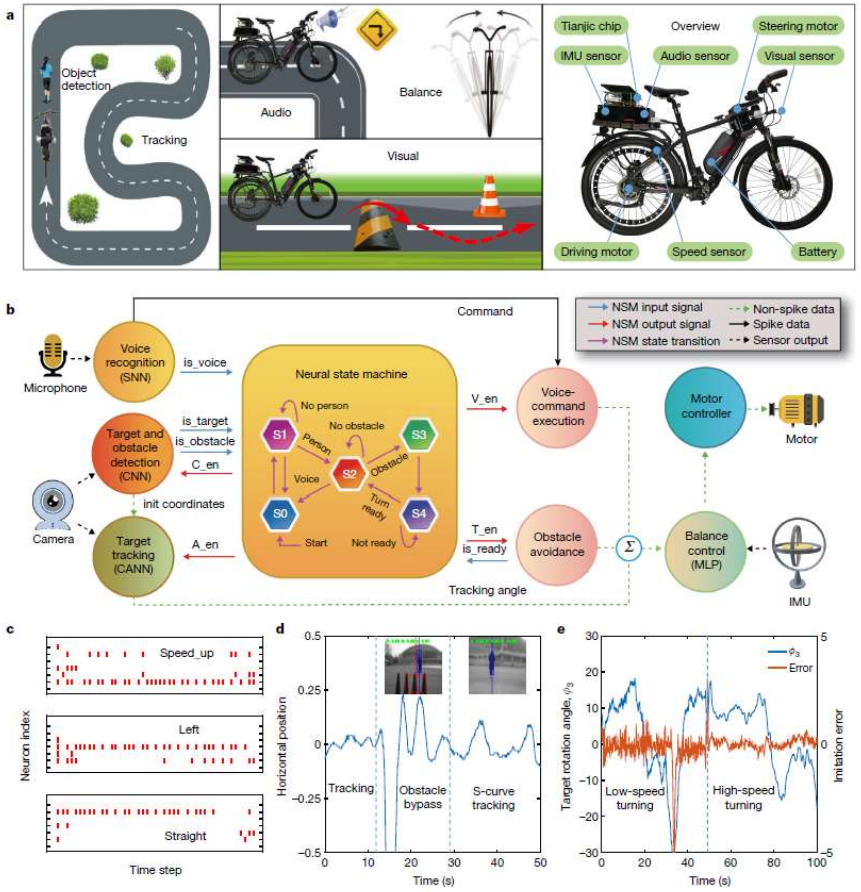


Fig. 4 | Demonstration of multimodal integration on a Tianjic chip for an unmanned bicycle. a, Left and centre, illustration of the tasks conducted in the bicycle experiment, including real-time object detection, tracking, voice perception, riding over a speed bump, automatic obstacle avoidance and balance of attitude. Right, the bicycle was equipped with a camera, gyroscope, speedometer, motors and a Tianjic chip. IMU, inertial measurement unit. **b**, Diagram of the multiple neural networks used in the unmanned bicycle experiment. The states inside the NSM diagram were defined as: voice command execution (S0), human detection (S1), human tracking (S2), a start of obstacle avoidance (S3), and a wait of avoidance completion (S4). An init coordinate is an initialization coordinate. C_en, A_en, V_en and T_en denote enabling signals for CNN, CANN, voice control and turning control, respectively. **c**, SNN voice-command-recognition test. The neuron producing the most spikes indicates the resulting classification. **d**, Tracking test. The *y* axis shows the relative horizontal position of the human in the frame. The bicycle automatically avoided an obstacle, then followed an instructor who ran in an S-curve route. **e**, Balance control and turning by an MLP network, which was trained by imitating the outputs of several well tuned controllers using the proportional-integral-derivative algorithm at different speeds (low to high).

Experimental Memrister based SNN (PIM)

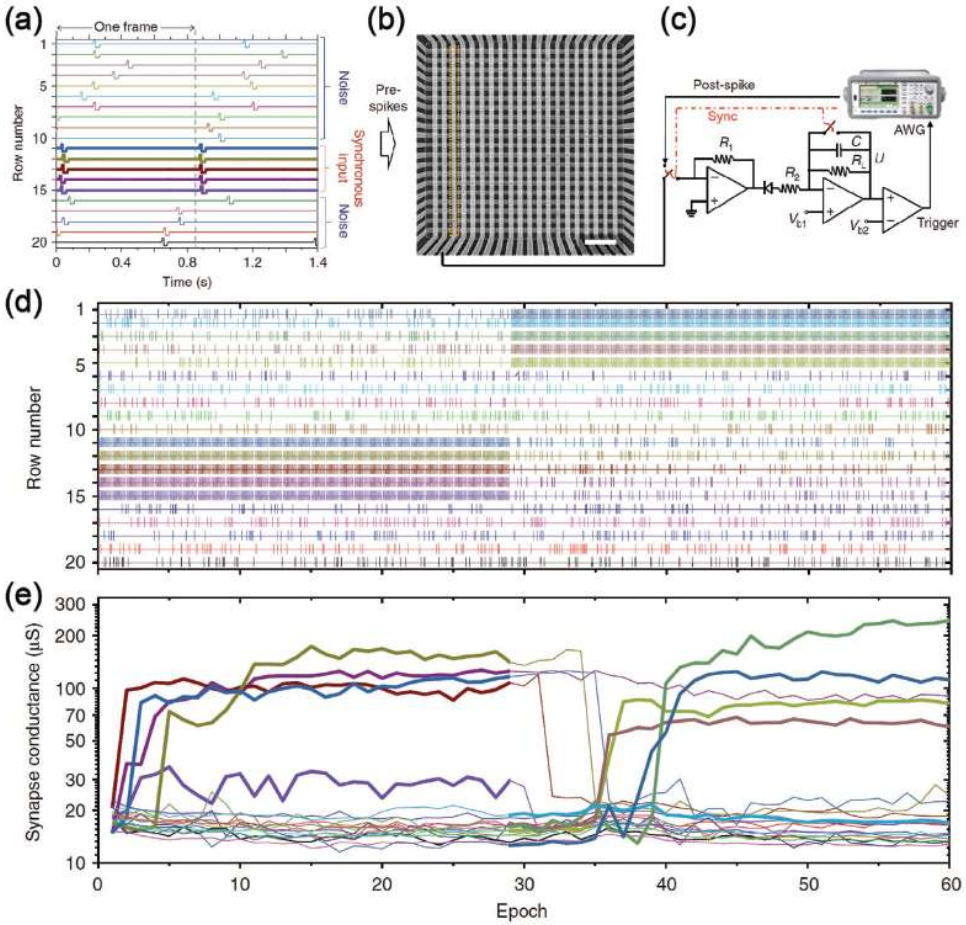


Figure 16. Demonstration of coincidence detection with a passive memristor-based SNN. a) An example of spike patterns applied to the synaptic array. b) Top-view of the scanning electron microscope image of the memristor crossbar array. Scale bar: 2 μm. c) Hardware implementation of the LIF neuron. d) Two lower-noise patterns for coincidence detection. The first correlated pattern was set to the rows ranging from 11 to 15 in the first 30 epochs, while the second pattern located in rows 1 through 5 in the following 30 epochs. e) Evolution in the conductance of all synaptic devices over time. a–e) Adapted under the terms of the CC-BY Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0).[214] Copyright 2018, The Authors, published by Springer Nature.

LIF and STDP are used to detect coincidence

SNN using RRAM based synapses

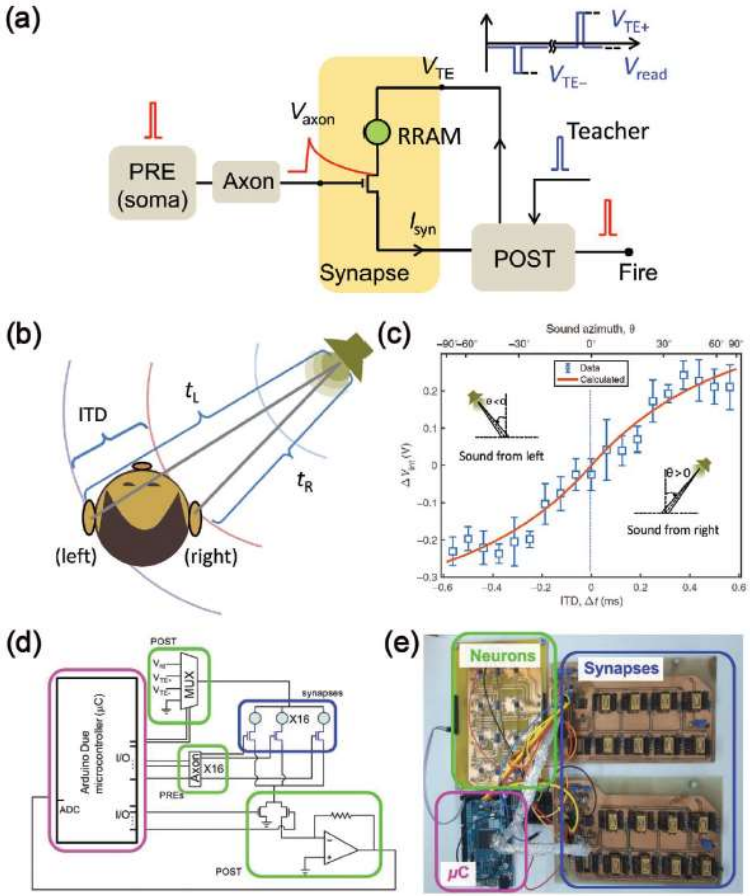


Figure 17. Learning of spatiotemporal patterns in a SNN. a) Schematic illustration of an artificial neuronal circuit consisted of a 1T1R synapse, a PRE axon, and a POST neuron which provides feedback signals for potentiating or depressing synaptic weights. b) Schematic diagram of sound location. Due to binaural effect, the ITD provides educible estimation about the position of sound with respect to the listener. c) Experimental results of sound location, revealing the relationship between ΔV_{int} and sound azimuth. d) Circuit connection diagram of the SNN, including synapses, PREs, a POST and a microcontroller. e) Photo image of the circuit configuration of the proposed SNN. a–e) Adapted with permission.[217] Copyright 2018, The Authors, published by American Association for the Advancement of Science (AAAS). Reprinted/adapted from ref. [217]. © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC) http://creativecommons.org/licenses/by-nc/4.0/.