Review
Competing memristors for brain-inspired computing

Seung Ju Kim,[1] Sang Bum Kim,[1,*] and Ho Won Jang[1,*]

SUMMARY

The expeditious development of information technology has led to the rise of artificial intelligence (AI). However, conventional computing systems are prone to volatility, high power consumption, and even delay between the processor and memory, which is referred to as the von Neumann bottleneck, in implementing AI. To address these issues, memristor-based neuromorphic computing systems inspired by the human brain have been proposed. A memristor can store numerous values by changing its resistance and emulate artificial synapses in brain-inspired computing. Here, we introduce six types of memristors classified according to their operation mechanisms: ionic migration, phase change, spin, ferroelectricity, intercalation, and ionic gating. We review how memristor-based neuromorphic computing can learn, infer, and even create, using various artificial neural networks. Finally, the challenges and perspectives in the competing memristor technology for neuromorphic computing systems are discussed.

INTRODUCTION

Unlike other living things, the greatest reason for mankind's prosperity is not physical strength but a highly developed intelligence. Human beings are thought to have the highest level of intelligence among intellectual creatures; this is presumed to be the result of having the most sophisticated brain. Based on these perceptions, human beings have been ultimately pursuing an understanding of the brain and intelligence. Since Turing pioneered artificial intelligence (AI), researchers have attempted to invent machines that can "learn" at a similar level as humans so that they can deal with the complex problems they faced. Consequently, the first programmable computers were invented beyond basic arithmetic operations. Turing suggested a conceptual machine that can solve any calculation using only proper rules (Turing, 1936). This conceptual architecture inspired the design of the von Neumann architecture, the foundation of modern computing architecture. The von Neumann architecture is a structure in which the processor and memory regions are physically separated, and data are transferred via bridged buses. Owing to the drastic advancement in complementary metal-oxide-semiconductor (CMOS) technology, computer performance has improved year by year, following Moore's law (Moore, 1965) and Dennard scaling (Dennard et al., 1974). Based on advances in computing performance and massive data explosions, recently developed AI algorithms based on deep learning have exhibited outstanding performance, especially in recognition problems.

However, big data produced by numerous edge devices and sensors require smaller, faster, and low-power-consuming computing performance, which has become a challenge with conventional CMOS-based computing systems. In addition, the high (energy and speed) cost of the von Neumann architecture has been touted as a problem, which is well known as the von Neumann bottleneck (Zidan et al., 2018). The latency between the processor and memory is inevitably generated in the von Neumann architecture because it separates the process units from memory units. To address these issues, researchers have returned their focus to the human brain and proposed brain-inspired neuromorphic computing.

Although the brain's learning mechanisms are not fully understood, the consistent development of neurobiology has revealed the operation principles of the brain, neurons, and synapses. Through a neurotransmitter, the synapses process and store information in a way that strengthens or weakens the connection between pre- and post-synaptic neurons, depending on the degree of involvement between neurons.

[1]Department of Materials Science and Engineering, Research Institute of Advanced Materials, Seoul National University, Seoul 08826, Republic of Korea

*Correspondence: sangbum.kim@snu.ac.kr (S.B.K.), hwjang@snu.ac.kr (H.W.J.)

Compared to the conventional CMOS-based computing that saves Boolean data type (0 or 1) in a single unit, the brain's neural networks save multiple states in a single synapse by adjusting the "synaptic weight," which is known as the synaptic plasticity. To apply these features to the new computing system, one memory cell needs to gradually change the information (Burr et al., 2017).

The "memristor" (portmanteau of "memory" and "resistor") technology, proposed by Chua (Chua, 1971), has been regarded as an emerging technology for realizing neuromorphic computing systems. Compared with the CMOS, the memristor exhibits an I-V hysteresis (Chua, 2014), which makes it possible to change conductance states gradually, similar to the biological synaptic weight. The memristor is also a nonvolatile memory technology that reads the nonvolatile resistive states, instead of the volatile capacitance states (Vourkas and Sirakoulis, 2015). In addition, these devices can be integrated into a large-scale crossbar array (CBA) architecture, where memristive devices are sandwiched between the crossed upper and lower electrode lines (Xia and Yang, 2019). Therefore, the theoretical density can be reduced to $4F^2$ (F is the feature size) and even higher with 3D structure (Lin et al., 2020), which is expected to overcome the scaling limit of the CMOS technology (typically exceeds $6F^2$).

In this review, we first introduce the six types of memristors (ionic migration, phase change, spin-based, ferroelectric, intercalation, and ionic gating devices) competing to be the representative next-generation neuromorphic device. We describe the operation principle and mechanism at the material nanoscale and illustrate how memristors act as artificial synapses in neuromorphic computing systems, similar to the human nervous system. Furthermore, we explore studies in which memristive neuromorphic computing has performed AI functions such as pattern recognition, classification, and even creation. Finally, perspectives and challenges in true brain-inspired computing are discussed.

COMPETING MEMRISTORS BASED ON DIFFERENT OPERATION PRINCIPLES

The concept of the memristor was first introduced as the fourth hidden fundamental element, in addition to the resistor, inductor, and capacitor. In 1971, Chua derived the mathematical equation $d\varphi = Mdq$, for finding a missing constitutive correlation between flux φ and charge q (Chua, 1971). The concept of a passive circuit element that satisfies such a correlation was theoretically needed, which Chua theorized and named "memristor," a portmanteau of memory and resistor. The theoretical verification of Chua's theory was confirmed by the invention of the first physical model of the memristor. In 2008, Hewlett-Packard researchers presented the first memristor devices based on a titanium dioxide insulator layer sandwiched between two metal electrodes (Strukov et al., 2008; Yang et al., 2008). This report provoked innovative electronic technologies and attempts to implement the memristor with various materials began.

Unlike existing CMOS-based memory technology, which reads volatile capacitance states, memristor technology is a nonvolatile technology that stores data using nonvolatile resistance states. All these devices exhibit an I-V hysteresis that changes from low resistance states (LRSs) to high resistance states (HRSs) and vice versa. Moreover, the area surrounded by the I-V hysteresis loop is continuously changed by changing the applied electrical signal. In other words, the memristor implements gradual conductance changes that can be utilized as synaptic plasticity.

Based on their operation mechanisms, memristors can be classified into six major types: ionic migration, phase change, spin-based, ferroelectric, intercalation, and ionic gating devices. These different operation principles determine the electrical behaviors of neuromorphic devices. Each device has been competing academically and commercially to become the key neuromorphic device in brain-inspired computing systems. In this section, these memristor mechanisms are described in detail.

Ionic migration

Ionic migration memristors were first introduced with TiO_2-based two-terminal devices developed by the Hewlett-Packard laboratory (Strukov et al., 2008). The ionic migration device is generally a two-terminal structure in which a solid-state dielectric material is sandwiched between the upper and lower metal electrode. Ionic migration memristors can also be classified according to the specific operating mechanisms through which the various types of carriers migrate in the dielectric layer between the electrodes. Most resistance changes, usually consisting of cations (such as active metal cations (Kim et al., 2019)), and anions (such as oxygen (Choi et al., 2013; Hsu et al., 2013; Jiang et al., 2016; Wei et al., 2008; Yoon et al., 2014, 2015),

halide (Choi et al., 2016; Han et al., 2019), nitride (Choi et al., 2012; Kim et al., 2017), and sulfurs (Xu et al., 2014; Zhang et al., 2016) ions) are induced, based on ion migration.

The cation migration mechanism mostly changes the conductivity by moving active metallic ions such as Ag and Cu under electric field control. In devices based on this mechanism, metal cations are separated from active electrodes and form a conductive filament through external electrical bias. A potential difference accelerates the migration of cations to achieve chemical equilibrium and leads to an electrochemical reaction. Consequently, the resistance state is changed from the LRS to the HRS when conducting bridges are formed between the top and bottom electrodes, resulting in a hysteresis loop in the positive voltage I-V characteristic. Because of this phenomenon, this type of mechanism is often referred to as "electrochemical metallization (ECM)."

In contrast to the set process, which is a process for forming conductive filaments, a reset process diffuses the filament and switches the resistance from the LRS to the HRS. A reverse electric field is applied to the rupture of filaments, which leads to a decrease in the conductance states and generates a hysteresis loop in the negative voltage I-V characteristic. Electron microscopy observations confirming the growth of Ag nanofilaments in SiO_2-based memristor (Cho et al., 2011; Sun et al., 2014) and voltage-activated rupture of filaments (La Barbera et al., 2015; Yang et al., 2012) support these observations.

Another major type of operation principle used for ionic migration memristors is the anion migration mechanism. The anion vacancies, counterparts of oxygen, halide, nitride, and sulfur ions, are utilized to achieve electron equilibrium by changing the valence state. As the anions migrate, the vacancies, which are their counterparts, migrate in the opposite direction. The anion vacancies inherent in the transition metal or related materials are more inclined to change the valence state of the cation under the electric field. The localized extended anion vacancies form defects and act as conductive filaments. Therefore, a stoichiometry change and a valence change of the cation leads to conductance changes between inert metal electrodes, such as Au and Pt. These processes are often collectively referred to as the "valence change mechanism" to distinguish them from the ECM.

Figure 1A presents the electrical characteristics and schematic operation of the ionic migration memristor. If a positive voltage is applied continuously, the conductive filament connects the upper electrode to the lower electrode and switches the HRS to the LRS in the set process. Conversely, when a continuous negative voltage is applied, the formed conductive filament ruptures and switches again from the LRS to the HRS through a reset process. Through this mechanism, the ionic migration memristor can induce a gradual conductance change and emulate the synaptic weight originating from the neural network.

Phase change

Since the phase change mechanism was first reported in 1968 (Ovshinsky, 1968), the phase change technology has been intensively studied as a candidate for nonvolatile memory technology; it has exhibited the potential for high-density integration on a very large scale. After it was reported that phase change memristors were fabricated using chalcogenide materials, such as $Ge_2Sb_2Te_5$ (Yamada et al., 1991) and Ag-and-In-doped Sb_2Te (Tominaga et al., 1997), many research groups intensively conducted related research to advance phase change memristors. Consequently, the phase change memristor technology is the most mature memory technology among competing memristor candidates; it has also been commercialized in storage class memory products using phase change technology (Burr et al., 2014; Hady et al., 2017). In addition, it has been suggested that this type of memristors also exhibits a gradual conductance change (Suri et al., 2012). Therefore, phase change memristors have been intensively studied for utilization in the neuromorphic devices (Tuma et al., 2016).

Figure 1B presents the electrical characteristics and principle operation schematic of the two-terminal phase change memristors. Similar to ionic migration memristors, phase change memristors also have the LRS and HRS. These devices operate based on the phase changes between a crystalline state and an amorphous state controlled by the difference in the applied heating time. In the crystalline state, long-range ordered vacancies with localized charge carriers cause the redesign of the band structure and increased conductivity (Pirovano et al., 2004; Siegrist et al., 2011). Accordingly, the crystalline phase exhibits LRS. An amorphous phase, on the other hand, has substantial structural randomness; therefore, it generates many localized electronic states that restrict electron mobility, resulting in HRS.
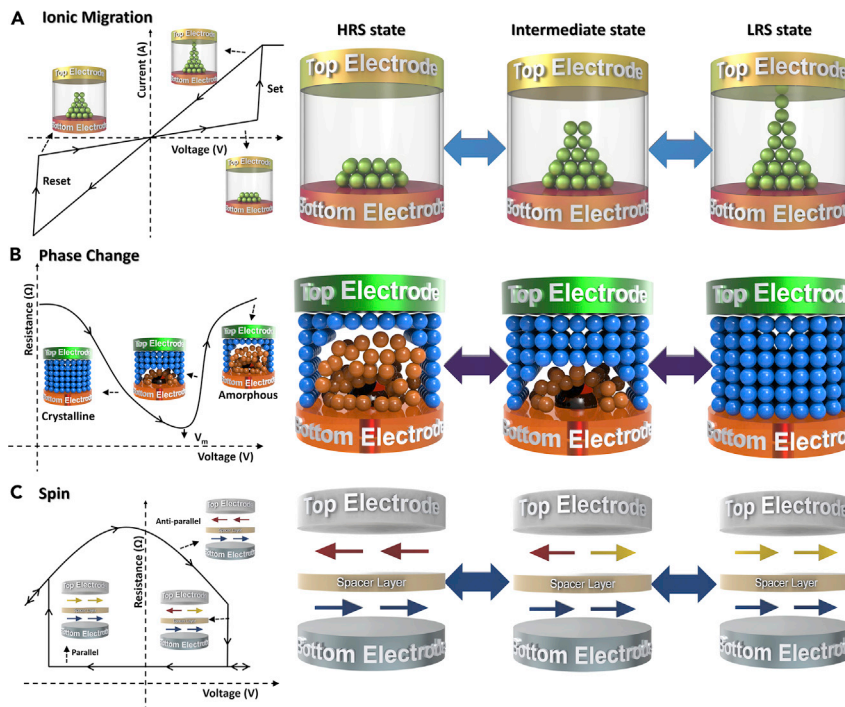
Figure 1. Memristive electrical characteristics and operation principle of ionic migration, phase change, and spin devices

Three types of memristors are presented. The electrical characteristics of each memristor exhibiting hysteresis: Schematics for typical current-voltage characteristics and operation principles with high resistance, intermediate, and low resistance states for (A) ionic migration, (B) phase changing, and (C) spin devices. All the bottom electrodes are grounded.

When applying a positive electric field between the top electrode and bottom electrode, which includes a narrow metal heater, a saturated current through the metal heater induces the Joule heating process. Continuously, the phase change material layer is heated above the melting temperature (generally $T_{melt} \approx 600°C$) using the metal heater's Joule heating. Following the rapid quenching of the under-cooled process, which usually takes sub-nanoseconds, the crystalline phase layer is switched to the amorphous phase. On the contrary, returning to the crystalline state involves maintaining the temperature of the amorphous state materials above the crystallization temperature (generally $T_{crystallize} \approx 400°C$) and slowly cooling it.

These mechanisms make it possible to control the gradual conductivity change during the set process, using unipolar programmed voltages. The conductance change is determined by the magnitude and length of the programmed voltage pulses applied to the metal heater. During the set process, applied low voltage pulses crystallize the phase change material, and the potentiated conductance increases gradually. Conversely, programmed voltages higher than the melting voltage form local melt regions that rapidly cool down to the amorphous phase, gradually decreasing the conductance.

Spin

Spin-based memristors have been studied since the giant magnetic resistance (GMR) phenomenon was first reported (Baibich et al., 1988). These devices utilize the magnetic hysteresis of ferromagnetic materials generated when an external magnetic field is applied. Owing to its low energy consumption, fast switching operation, and excellent precision, the spin memristor has been regarded as an excellent potential candidate for neuromorphic computing systems (Grollier et al., 2016, 2020; Locatelli et al., 2014). Figure 1C presents a schematic diagram of the resistive voltage hysteresis loop and spin-based memristor principle operation process.

A magnetic tunnel junction (MTJ) with a spin-valve structure is a basic unit of the spin-based memristor. In the MTJ unit, a non-magnetic metal layer (described as a spacer layer) is laminated between two

ferromagnetic layers. The early spin memristor used the GMR phenomenon to adjust the conductivity. The relative magnetization direction of the ferromagnetic layers is switched by an interlayer exchange coupling induced by the external magnetic field (Dieny et al., 1991). If the magnetization direction of the adjacent ferromagnetic layers is anti-parallel, electrons with different spin orientations are strongly disturbed by spin-dependent scattering and reach the HRS. Conversely, when the relative magnetization is parallel, the flow of electrons between the ferromagnetic layers becomes easier, leading to the LRS.

Because the conventional GMR technology manipulates the spin states by applying external magnetic fields, inevitable magnetic interference occurs, as the distance between adjacent MTJs decreases. In 1996, Slonczewski first reported the spin-transfer torque (STT) effect that can manipulate the spin state without any external magnetic field (Slonczewski, 1996). He proposed that the spin-polarized current flowing through an MTJ adjusts the alignment of the spin orientation using the tunneling magnetoresistance (TMR) phenomenon. An MTJ unit consists of a thin spacer layer functioning as a tunnel barrier and two metallic ferromagnetic layers. A thicker ferromagnetic layer acts as a pinned layer (PL) because its spin polarization is pinned in a certain direction. The other thinner ferromagnetic electrode layer acts as a free layer (FL) that can modulate the spin direction through the STT effect of the polarized current. When electrons flow through the pinned layer, their spin orientation conforms to the same orientation as the PL. These spin-polarized electrons are partly sent through the tunnel barrier to transfer the spin angular momentum to the FL. Subsequently, the spin orientation of the FL is switched to the PL, resulting in a parallel configuration and decreasing resistance. If the current direction is inverted, the electrons with opposite spin polarization to the FL are reflected by the PL back to the FL, and MTJ cells are switched to an anti-parallel state. In summary, the spin-based memristors typically require a bipolar operation because the direction of the current flowing between the MTJ cell determines the spin orientation of the FL. In addition, the spin-orbit torque effect (Brataas and Hals, 2014; Liu et al., 2012), spin hall effect (Zhang, 2000), antiferromagnetism (Železný et al., 2018), and skyrmion (Huang et al., 2017; Song et al., 2020) can be used in spin-based neuromorphic systems.
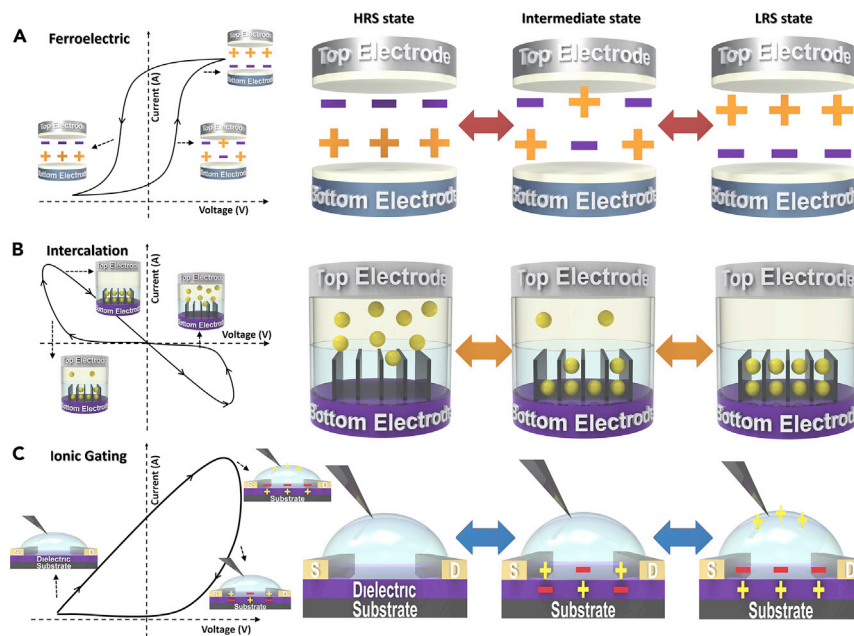
Ferroelectricity

Ferroelectricity is an electrical characteristic of certain materials that have spontaneous electric polarization induced by an external electric field. This phenomenon was initially used for nonvolatile memory applications, such as ferroelectric random-access memories (Scott and Paz De Araujo, 1989). However, the ferroelectric film growth technology was limited in terms of nanoscale fabrication; thus, the ferroelectric memristor did not attract attention. Recently, it was reported that ferroelectric materials based on doped HfO_2 can be fabricated as few nanometer layers (Böscke et al., 2011; Mueller et al., 2012; Müller et al., 2011) through the atomic layer deposition process, and the interest in ferroelectric memristors has been revived. The concept of the ferroelectric tunnel junction (FTJ) was introduced (Chanthbouala et al., 2012; Tsymbal and Kohlstedt, 2006), and remarkable progress has been reported for neuromorphic applications (Boyn et al., 2017; Li et al., 2020; Luo et al., 2019; Majumdar et al., 2019).

Figure 2A shows a schematic diagram of the I-V characteristics and operating principles of ferroelectric memristors. A typical FTJ is a two-terminal metal/ferroelectric film/metal structure. An ultra-thin tunnel barrier sandwiched between two different metal electrodes forms two different interfaces. These interfaces assume an asymmetric form for the energy band profile. When an electric field is applied, the uncompensated polarization charges accumulated on the interfaces change the average barrier height of the electric potential and attract or repel electrons, based on their sign. In other words, the uncompensated polarization charges partially screen the free carriers. Although the screening length is predicted by Thomas-Fermi theory, which states that the theoretical screening length is determined by the density of states at the Fermi level, the effective screening length strongly depends on the microscopic properties at the interfaces (Stengel et al., 2009).

The interface thickness between the metal and ferroelectric material is affected by the alignment of the ferroelectric polarization and manipulates the asymmetry of the potential barrier. Subsequently, the tunneling current between the ferroelectric memristor can be controlled by the alignment of the polarization in the ultra-thin ferroelectric layer. This screening phenomenon is called the tunneling electro-resistance (TER) effect (Garcia et al., 2009). The TER ratio is defined as follows:

$$\text{TER ratio (\%)} = \frac{R_{P_+} - R_{P_-}}{R_{P_-}} \times 100$$

Figure 2. Memristive electrical characteristics and principle operation of ferroelectric, intercalation, and ionic gating devices

Three types of memristors are presented. The electrical characteristics of each memristor exhibiting hysteresis: Schematics for typical current-voltage characteristics and operation principles with high resistance, intermediate, and low resistance states for (A) ferroelectric, (B) intercalation, and (C) ionic gating devices. All the bottom electrodes are grounded.

Ferroelectric memristors use this phenomenon to gradually control the conductivity in neuromorphic computing systems. Furthermore, ferroelectric field-effect transistors (FeFETs) have been actively researched to be used in ferroelectric-based neuromorphic computing systems. By manipulating the polarization of the ferroelectric layer, FeFETs can operate at extremely low power consumption (Dünkel et al., 2018; Oh et al., 2019).

Intercalation

The lithium-ion intercalation effect has been widely studied in the anode materials of battery cells (Sasaki et al., 2013). The principle of lithium-ion secondary batteries is as follows. During the charging process, the lithium ions separated from the cathode pass through the electrolyte and separator based on the applied electrical potential. Subsequently, the lithium ions intercalate between the layered anode materials, such as graphene nanosheets, and the accumulated electrical potential is stored, even after the external voltage is turned off. In the discharge process, the lithium which is intercalated in the layered anode materials is oxidized by the loss of electrons and deintercalated from the anode materials. The separated electrons flow from the anode to cathode along the external line and reduce the cathode materials such as $Li_{1-x}CoO_2$. Because this charge-discharge process is fully reproducible and sufficiently repeatable, the preference for electrochemical technology has grown dramatically in the battery market. Ever since the memristive phenomenon of the Li_xCoO_2 thin layer that caused by the cobalt redox reaction coupled to the lithium-ion intercalation in the conducting silicon electrode was reported (Moradpour et al., 2011), various attempts have been made to study the use of lithium-ion intercalation for implementing neuromorphic devices (Choi et al., 2019; Ioannou et al., 2020; Mai et al., 2015).

Unlike other memristors, the intercalated lithium ions remain in the electrolyte, even after the applied external voltage is removed. Furthermore, the concentration of lithium ions in the electrolyte can change the conductivity linearly, which cause the I-V hysteresis loop. This provides a relatively constant memory effect in the intercalation memristors.

Figure 2B shows the I-V characteristics and operation mechanism principle of the intercalation memristors. If an external electric field is applied to the cathode (described as a top electrode), atoms in the lithium

metal cathode are positively ionized by the oxidation reaction. The lithium ions then migrate toward the layer-shaped anode (described as a bottom electrode) and are intercalated into the anode layer through the external electric bias. Because the intercalation of lithium ions changes the electronic distribution surrounding the lithium ions, the conductance of the intercalation-based memristor increases. The number of ions driven into the anode layer can increase the conductance of devices. Conversely, when a reverse bias is applied to the cathode, the lithium ions are deintercalated from the anode layers and return to the cathode. Consequently, the conductance of the memristors decreases.

The electron band structure has demonstrated the intercalation mechanism for neuromorphic devices. This mechanism is related to the electron distribution surrounding the intercalated ions. The migrated lithium ions modify the electron distribution near the lithium ions and form a local conduction region. The movable ions physically change the band structure of the devices, resulting in a change in conductivity.

Moreover, three-terminal synaptic transistors, which have channels and electrolytes, have been actively researched, for use in neuromorphic computing systems (Fuller et al., 2017; Sharbati et al., 2018; Yang et al., 2018a; Zhu et al., 2018). The lithium ions in electrolytes are intercalated into the channel material between the source and drain electrodes. Because these devices have relatively superior conductance change linearity, compared to other memristors, various configurations of channel materials and electrolytes have been intensively studied.

Ionic gating

The most striking difference of ionic gating memristors is that it has an ionic liquid or ionic electrolyte gate. For this reason, it has mainly been used for synaptic transistors with a three-terminal structure. Transistors with ionic gel/liquid electrolytes have been studied for a long time before attention shifted to neuromorphic devices. Because these devices operate at relatively low voltages, their energy consumption is relatively low. In addition, the change in ionic charge memristor conductivity is considerably more linear than in other memristors, which encourages attempts to utilize them in neuromorphic computing systems. The first synaptic transistor based on an ionic liquid and $SmNiO_3$ was implemented in 2013 (Shi et al., 2013). This memristor exhibited an extraordinarily linear conductance change, compared to conventional memristors. Subsequently, ionic memristors with diverse material configurations, such as α-MoO_3/ionic liquid (Yang et al., 2017), WO_3/ionic liquid (Yang et al., 2018b), SFO/ionic liquid (Ge et al., 2019), and ETE-S/ionic liquid (Gerasimov et al., 2019), have been proposed.

Figure 2C is a diagram of the I-V characteristics and operation mechanisms of ionic gating memristors. In ionic gating memristors, the concentration of protons or oxygen ions generally causes a gradual change in conductance when an external electrical bias is applied to the ionic electrolyte gate. An external negative voltage is applied to the ionic electrolyte gate, generating reduced charge carriers in the gate. Subsequently, the electrical potential causes a partial distribution of the charged carriers. The charges gathered around the channel materials induce an opposite charge to the surface of the channel. This has an effect similar to doping the surface of the channel material, increasing its conductivity. Conversely, when the reverse voltage is applied, the doping concentration in the channel material is lowered, and consequently, the conductivity of the memristor decreases. Following this process, the charge carrier adjusts the conductivity by doping the channel material with an electric charge. Although ionic gating memristors have only recently been proposed for neuromorphic devices, they have been rapidly developing. In the next chapter, we will demonstrate how these memristors act as building blocks for neuromorphic computing systems and how they compete with each other to implement synaptic characteristics.

MEMRISTORS IN NEUROMORPHIC COMPUTING SYSTEMS
Artificial synapses for neuromorphic computing

The nervous system of vertebrates (including human beings) consists of the brain, spinal cord, and peripheral nerves. The brain is an organ that controls the central nervous system and is known as the center of intelligence. It receives sensory information from peripheral nerves all over the body and coordinates the actions necessary for life activities. In the human brain, which is considered the most advanced, computations are performed to support intelligent functions such as memorization, forgetting, and decision-making, which are executed through the interaction of neurons in the brain. Therefore, humans can perform advanced intellectual activities such as learning. Neuromorphic technology was inspired by the brain's mechanisms and has been developed by mimicking the brain composition at each level. Figure 3
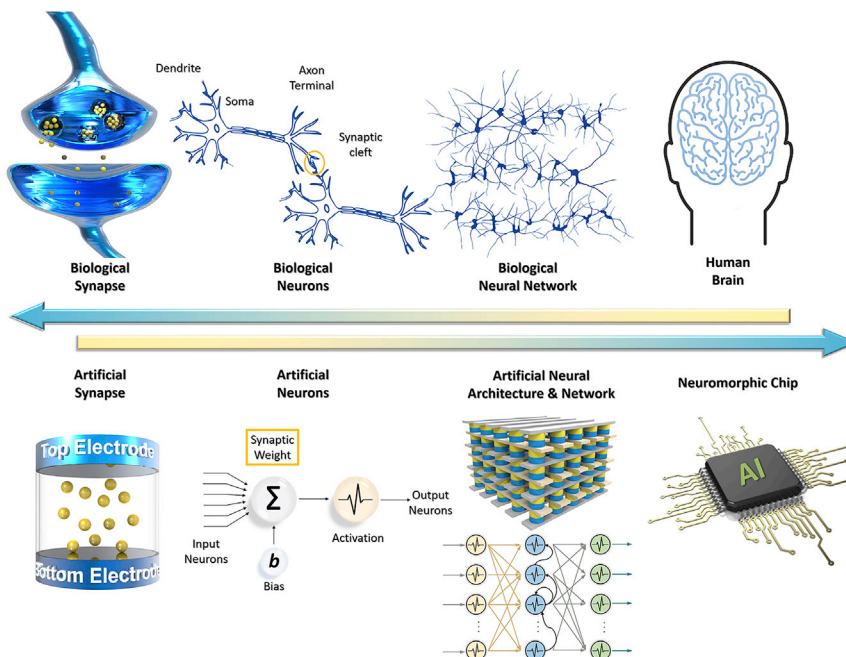
Figure 3. In-depth comparison of human nervous system and artificial neural system in neuromorphic devices

presents a side-by-side detailed comparison between the human nervous system and an artificial neural system in neuromorphic devices.

In the human nervous system, the human brain performs various computational functions, such as recognition, memorization, and even creation, through the activities of numerous biological neural networks. These neural networks are composed of the complex connections of neurons. A typical neuron consists of the soma, which is a cell body, many dendrites that receive input signals, and axons with many terminal branches that transfer output signals to other neurons. The axon terminal connects to the dendrite of another neuron, and this specialized connection is called a synapse. The adult's brain is estimated to consist of approximately 8.6×10^{10} (86 billion) neurons (Azevedo et al., 2009) and 5×10^{14} (500 trillion) synapses (Drachman, 2005). Different types of ion channels in the synaptic cleft generate the molecular basis of electrical activity for the target neuron. Through the type and number of neurotransmitters, the synapses can be excited or inhibited to support information processing. In other words, learning is achieved through the ability of synapses to reconfigure their strength connecting neurons. This ability is called synaptic plasticity, which has long been of interest in the neuroscience field.

Focusing on these features, a neuromorphic chip artificially emulates different parts of the biological nervous system. Various types of artificial neural network (ANN) algorithms implement various computational functions for neuromorphic chips. These algorithms process the calculation in the architecture of the CBA structure and neural circuits. In the CBA, the word line (WL) electrodes cross over the bit line electrodes, and the selector (S) screening leakage current and memristor (R) are placed at all the cross-points between each electrode to act as a single memory cell (1S-1R). Unlike the conventional CMOS-based von Neumann architecture, the CBA has the advantage of parallel accessibility during operation, saving multiple levels in one memory cell instead of Boolean values, and reducing the scaling limit to $4F^2$ or even lower. An artificial neuron that is mathematically defined (e. g., sigmoid or rectified linear unit) or composed of a neural circuit (such as a leaky integrate-and-fire model) is activated when the synaptic weight value stored in each memory cell is given, following which it updates information through mathematical or time-dependent feedback. After operations, the computed results are converted to programmed voltage signals and apply them to the WL to change the synaptic plasticity in a memristor. Synaptic plasticity can be implemented by gradually modifying the conductance. As aforementioned, it is implemented using ionic migration, phase change, spin, ferroelectric, intercalation, and ionic gating operation mechanisms. Therefore, in neuromorphic computing systems, these memristors act as artificial synapses, which are the building blocks of the artificial neuron, network, and intelligence chip technology.
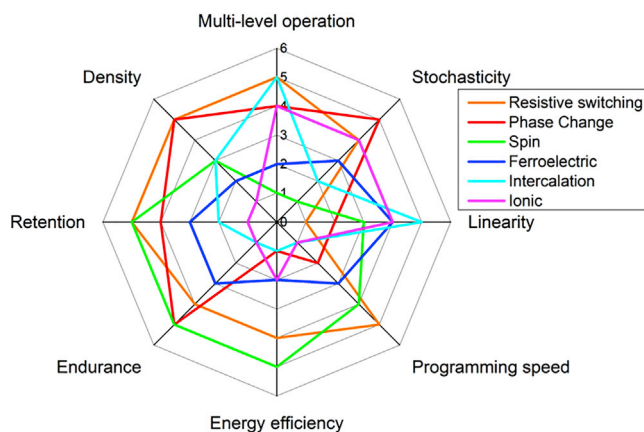
Figure 4. Comparison chart of neuromorphic characteristics of memristors used as brain-inspired computing systems

Artificial neurons for neuromorphic computing

Apart from advances in artificial synapses, hardware implementation of artificial neurons using memristor has remarkably progressed in brain-inspired computing technology. The basic features of artificial neurons are that they can accumulate signals and fire spiking when the stimulus intensity reaches a threshold. In biological neurons, sensory or chemical stimuli cause changes in synaptic potential. The soma integrated these activities and determines the axon potential. If the accumulated stimulus exceeds the threshold potential, the same output signal occurs no matter how much the input potential exceeds; otherwise, there is no output response. This characteristic follows the "all-or-nothing" or "all-or-none" law.

To reproduce the integration and firing characteristic of biological neurons, abundant research has been carried out with emerging memristor technology. Unlike mathematical defined artificial neurons, as mentioned in section 3.1, memristor devices itself can emulate the basic function of biological neurons such as integration and fire function.

Ionic migration memristor was used to demonstrate the integrating and firing function of artificial neurons. The non-volatile threshold switching phenomenon of ion-based (included anion and cation) conducting filament can emulate the accumulation process of membrane threshold potential (Mehonic and Kenyon, 2016; Wang et al., 2018a; Woo et al., 2017). Similarly, the phase change memristor has been proposed to mimic the stochastic firing response of neurons by the phase transition of the chalcogenide film (Tuma et al., 2016). Furthermore, spin-based memristor (Fan et al., 2015; Sharad et al., 2013) and ferroelectric field-effect memristor (Mulaosmanovic et al., 2018) have been demonstrated to emulate integrated-and-fire characteristics. Intercalation and ionic gating memristors still have been researched to implement artificial neuron characteristics in many research groups.

Characteristics of competing memristors for neuromorphic technology

Advanced CMOS technology facilitates the execution of advanced AI algorithms with conventional von Neumann computing systems. However, von-Neuman computing has reached its limit because of the speed delay, arising from the von Neumann bottleneck, huge energy consumption, and an inherent architecture that can store only Boolean data (0 and 1). To address these problems using memristors, it is necessary to satisfy the characteristics of neuromorphic computing and achieve excellent performance, compared to existing devices. The required characteristics are as follows: ability to express synaptic weights and plasticity; fast operations; low energy consumption; and the ability to store information reliably and integrate it in high density. Memristors with different operating mechanisms are competing to satisfy these characteristics. Figure 4 and Table 1 compare the state-of-the-art performance of six types of memristors in the implementation of neuromorphic computing systems based on the eight requisite characteristics (Tang et al., 2019; Wang et al., 2020; Xia and Yang, 2019).

	Ionic migration	Phase change	Spin	Ferroelectric	Intercalation	Ionic gating
Multi-level operation	High	Moderate	Low	Moderate	High	Moderate
Stochasticity	Moderate	High	Low	Moderate	Moderate	Moderate
Linearity	Low	Moderate	Moderate	Moderate	High	High
Programming speed	High	Moderate	Moderate	Moderate	Low	Low
Energy efficiency	Moderate	Moderate	High	Low	Moderate	Moderate
Endurance	Moderate	High	High	Moderate	Low	Low
Retention	Moderate	High	High	Moderate	Moderate	Low
Density	High	High	Moderate	Moderate	Moderate	Low

Table 1. Comparison of neuromorphic characteristics of memristors to be used as brain-inspired computing systems

Synaptic plasticity is a key concept in neuroscience and brain-inspired computing systems. As aforementioned, synaptic plasticity is the ability of neurons or synapses to change their synaptic weight. Typically, in electronic devices, the synaptic weight is expressed by changing the device conductance. To express the synaptic weight and plasticity in electronic devices, multi-level states in conductance, stochasticity, and linearity are required.

Tunable multi-states operated by potentiation and depression provide a huge storage capacity, compared to digital devices, which save only two types of information: 0 and 1. Ionic migration and intercalation memristors exhibit good multi-level operations because their ions migrate gradually in a dielectric layer. Stochasticity indicates how much conductance fluctuates randomly over a given period of time. Unlike deterministic digital computing, random noise in neuromorphic computing helps with learning and information processing. Phase change exhibits good stochasticity because of the reconfiguration induced by the melting and quenching in the atomic structure of the phase change layer (Tuma et al., 2016). Linearity indicates how uniformly the conductance increases and decreases, as the weight update pulse number increases. The conducting filaments of ionic migration memristors are repeatedly cut off, which results in low linearity. On the contrary, intercalation and ionic gating memristors exhibit excellent linearity and symmetry because their conductance states are adjusted by their concentration changes.

The programming speed and energy efficiency determine the performance of the neuromorphic computing that overcomes the von Neumann bottleneck. Phase change memristors perform poorly because they require a large current to elevate Joule heating, and a relatively long time is required for crystallization. Owing to nanoscale channels and quantum-scale physics characteristics, the ionic migration and spin memristors exhibit outstanding performances.

To ensure the device's reliability, high endurance and retention are required. Although these characteristics are typically related to nonvolatile memory technology, they are also required in neuromorphic computing systems. Endurance is the maximum number of switching cycles between on and off. Retention indicates how long the stored data are maintained. The spin and phase change memristors are mature technologies in the nonvolatile memory field, and they exhibit high endurance and retention.

The more densely integrated devices can perform in smaller neuromorphic computing systems. The density indicates the subarray size that can be fabricated. However, the more densely integrated into crossbar array architecture, the more leakage current through the unselected cells, which is known as the "sneak current" issue. To address this issue, the operation voltage schemes and selection devices are proposed (Im et al., 2020). The ionic migration and phase-change memristors can be fabricated on a sub-nanometer scale. On the other hand, the ionic memristor is composed of a gel or liquid and should have a three-terminal structure, which results in poor density. Furthermore, the intercalation and ionic gating memristors are relatively less compatible with the CMOS fabrication processes. For example, electrolyte and electrode materials for the intercalation memristors are usually not stable in air, and the ionic gating materials are not sufficiently rigid to yield well-defined nanostructures. Including these eight types of synaptic characteristics, various features of biological neural characteristics, such as diffusive dynamics (more detail in section 4.3),

have been attempted to implement in electronic hardware and software levels. In summary, each memristor technology has been competing to achieve the requisite characteristics of neuromorphic computing systems.

In the next chapter, we will explore how these memristors are utilized by the latest ANN algorithms to solve a variety of problems, including recognition, classification, and creation.

ARTIFICIAL NEURAL NETWORKS WITH MEMRISTOR

Since the first mathematical model of ANNs was proposed by McCulloch-Pitts in 1943 (McCulloch and Pitts, 1943), various attempts have been made to implement human nervous systems as mathematical algorithms. Perceptron (Rosenblatt, 1958), the first algorithm implemented in 1958 by Rosenblatt, was not significantly remarkable, due to insufficient computational performance. Over the past few decades, the CMOS technology has improved rapidly, following Moore's law. Subsequently, significant progress in ANNs has contributed to progress in AI applications, including speech recognition (Mohamed et al., 2012; Xiong et al., 2017; Young et al., 2018), face recognition (He et al., 2015; Taigman et al., 2014), and medical prescription (Jude Hemanth et al., 2014; Litjens et al., 2017).

ANNs are network algorithms that have layers and activation functions with a weighted sum of inputs to output values. The neurons and synapse models used in current ANNs are simple, compared to human nervous systems. ANNs are usually classified based on the main synaptic plasticity used and the connectivity structure, such as convolutional neural networks (CNNs), deep neural networks (DNNs), long short-term memory (LSTM), spiking neural networks (SNNs), and generative adversarial networks (GANs). They make it possible for computing systems to perform various computational functions such as recognition, memorization, and creation. While conventional CMOS-based von Neumann computing systems implement these necessary synapse functions in numerous sets of Booleans, memristor-based neuromorphic computing systems allow a single memristor unit to implement a single artificial synapse. In this chapter, we will demonstrate how memristors can implement a neural network algorithm using synaptic plasticity.
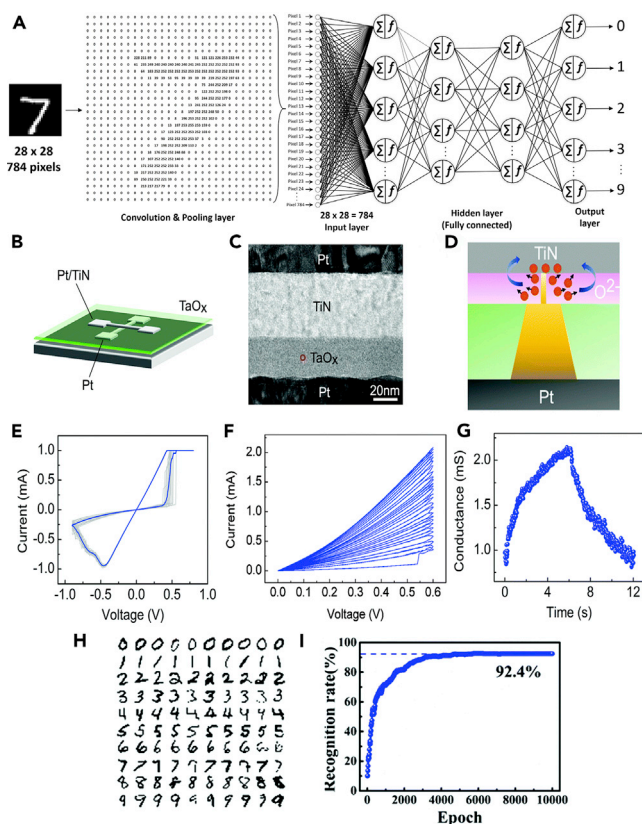
Convolution and deep neural networks with memristors

CNNs are commonly used networks for extracting features in visual image analysis (LeCun et al., 1998). The CNN computes a convolution operation to map extracted features from given input images and perform pooling operations to store only the representative values. CNNs consist of mathematically defined layers that operate algorithms, and neural networks consist of more than three layers, which include multiple hidden layers, and are known as DNNs (Sze et al., 2017). These layers typically include convolutional layers, pooling layers, activation functions layer, and hidden layers with fully connected layers.

Figure 5A illustrates the general procedure of implementing the CNN and DNN with the modified National Institute of Standards and Technology (MNIST) handwritten number recognition. Their learning process is as follows: Features are extracted from a given input image through the convolution layer, following which the representative values are resized, and stored in the pooling layer. Subsequently, flattened $28 \times 28 = 784$ data are input into an input layer to compute the synaptic connectivity between the fully connected layers. These networks are trained by supervised or unsupervised learning through backpropagation operations, including stochastic gradient descent (Ruder, 2016) and chain rules (Caterini and Chang, 2018). Consequently, they optimize the best-deduced model by adjusting the synaptic weight. The presence of hidden layers certainly increases the recognition ability significantly, comparable to the human brain.

However, the inevitable flaw of these algorithms is that the complexity of computation (such as convolution, full connection, and backpropagation) increases the computational cost. Conventionally, the CMOS-based CPUs have implemented CNNs and DNNs with the help of accelerators, including graphic processor units, field-programmable gate arrays, and application-specific integrated circuits that are designed for very parallel computation. However, they are still prone to high power consumption and inefficient processing of synaptic weight in bits.

Memristor technology can implement mathematical synaptic weights in the algorithm by gradually changing the conductance of a memristor unit. This effectively improves the performance of existing CNNs and DNNs. Synaptic weight changes can be achieved by gradually increasing and decreasing the conductance in memristors. Synaptic weight changes are also referred to as long-term potentiation and long-term

Figure 5. Convolution and deep neural networks with memristors

(A) Schematic of multi-layer convolution and deep neural networks with MNIST pattern recognition.

(B–G) (B) Device schematic, (C) cross-sectional TEM image, (D) operation principle diagram, (E) current-voltage characteristic, (F) analog potentiation switching of TaOx ionic migration memristor, and (G) potentiation and depression of conductance. Reproduced with permission (Wang et al., 2016). Copyright 2016, Royal Society of Chemistry.

(H and I) (H) MNIST data set for pattern recognition deep learning and (I) accuracy of recognition by iteration. Reproduced with permission (Wang et al., 2019). Copyright 2019, Royal society of Chemistry.

depression, which are induced by applying programmed voltage pulses and maintaining the conductance state for a long time.

Wang et al. reported a TaO$_x$-based ionic migration memristor for neuromorphic devices (Wang et al., 2016). Figure 5B shows a schematic of the TiN/TaO$_x$/Pt memristor as a 2-terminal structure. As confirmed in the cross-sectional TEM image in Figure 5C, the TaO$_x$ layer, approximately 25-nm thick, was deposited by the radio frequency (RF) sputtering. The TiN top electrode and Pt bottom electrode act as the pre- and post-synaptic terminal, respectively. The oxygen vacancies in TaO$_x$ form a conducting filament in the memristor, which is described in Figure 5D. Therefore, the I-V characteristic shows the pinched hysteresis loop in Figure 5E, which corresponds to memristive behavior. In addition, the continuous direct current (DC) sweeps with 27 modulation steps affect the conductance modulation. Based on these characteristics, the potentiation and depression synaptic plasticity are implemented by applying 300 identical 1V/100ns positive voltage pulses and 300 identical −1.1V/100ns negative voltage pulses. In addition, a variety of memristors have been observed to exhibit these synaptic potentiation and depression characteristics, including ionic migration (Jo et al., 2010; Li et al., 2013), phase change (Suri et al., 2012), spin (Wang et al., 2018b), ferroelectric (Jerry et al., 2017), intercalation (Choi et al., 2019; Fuller et al., 2017; Ioannou et al., 2020; Yang et al., 2018a), and ionic gating (Shi et al., 2013) memristors.

Using this synaptic plasticity, a neural network can effectively perform pattern recognition with memristor devices. Wang et al. successfully fabricated a fully transparent, flexible, and waterproof organic memristor (Wang et al., 2019). They presented a possibility to perform MNIST handwriting recognition using the
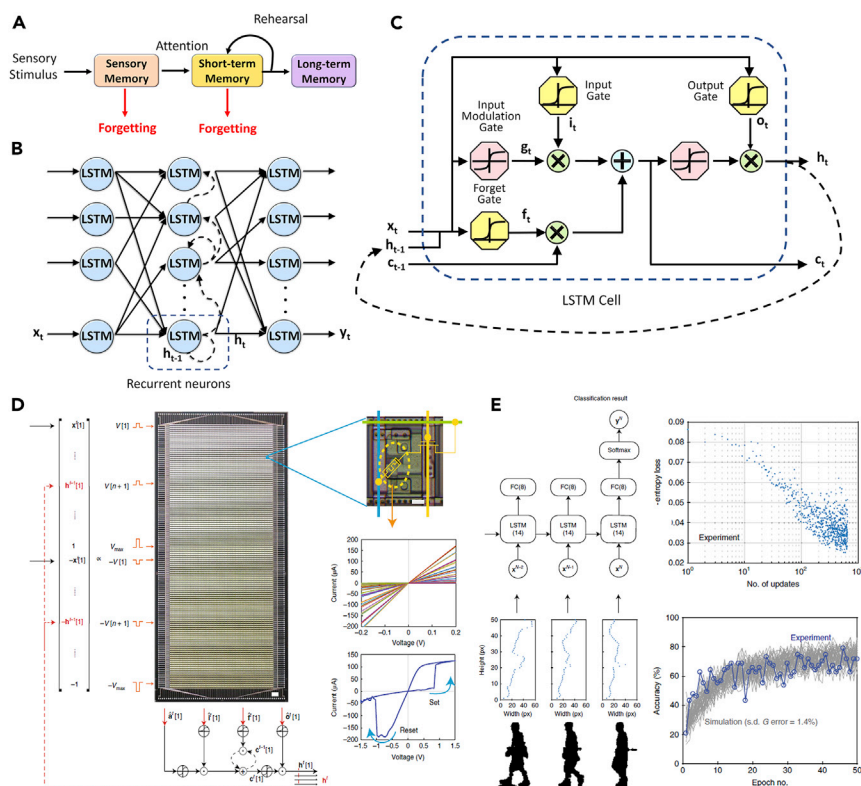
Figure 6. Long short-term memory networks with memristors
(A) Multi-store model of memory mechanism proposed by Atkinson-Shiffrin.
(B) Schematic drawing of multi-layer recurrent neural networks with long short-term memory nodes.
(C) Architecture of the LSTM cell with logcial algorithms.
(D) 128 × 64 1T-1R memristor CBA with LSTM units.
(E) Accuracy of human recognition by gait.
Reproduced with permission (Li et al., 2019) Copyright 2019, Springer Nature.

circuit simulation. Based on these synapses, the trained virtual hardware can perform near the theoretical limit. Figure 5H is a sample image of the MNIST handwritten numbers. These images were input to the neural network algorithm, which is composed of an input layer with 256 neurons, a hidden layer with 128 neurons, and an output layer with ten neurons. Through supervised learning, the recognition accuracy was approximately 92.4% after 5000 epochs. Furthermore, the recently developed actual neuromorphic CBA chips have demonstrated outstanding performances to implement AI applications (Ambrogio et al., 2018; Cai et al., 2019; Li et al., 2018; Yao et al., 2020).

Recurrent and long short-term memory network with memristors

A human memory process model was first proposed in 1968 by Atkinson and Shiffrin (Atkinson and Shiffrin, 1968). This model divides the human memory into sensory memory, through which sensory stimulus enters memory, short-term memory that saves information temporarily, and long-term memory, where information is indefinitely stored through rehearsal in the short-term memory. Figure 6A describes the multi-storage model of human memory. Based on this concept, a recurring neural network (RNN) that can process sequential information has been developed.

Unlike other neural networks, the RNNs have different types of connections between neurons in the same layer. In the RNN architectures, the neurons connect to each other or themselves, enabling information to flow in a loop, which enables the implementation of sequential dynamic features. This architecture allows the RNN to process data with sequences such as natural language processing and speech recognition. However, as the gap in order grows in neuron processing, RNNs miss the sequential connection between

data, which is called the "gradient vanishing problem". This problem, known as "long-term dependency," was explored in depth by Bengio (Bengio et al., 1994).

In 1997, Hochreiter and Schmidhuber proposed LSTM networks to address this issue (Hochreiter and Schmidhuber, 1997). LSTM networks are based on RNN and are capable of learning long-term dependencies. Figures 6B and 6C illustrate the schematic diagrams of the LSTM network based on the RNN and LSTM neuron structure. Generally, the LSTM cell consists of an input gate, an input modulation, a forget gate, and an output gate. The new input x_t and previous output h_{t-1} pass through a sigmoid function of the forget gate where it is determined whether they will be used or not. Output 1 indicates that the value is maintained, and 0 indicates a complete deletion. If the output of the forget gate is 1, the sigmoid and tanh functions of the input layer generate values and update the internal memory from C_{t-1} to C_t. The output gate refers to the internal memory and calculates the output h_t. This recurrent processing based on the internal memory value with the backpropagation algorithm solves the gradient vanishing issues and significantly improves the effectiveness of the system. However, it is inevitable to compute large amounts of backpropagation and even internal memory data to perform the LSTM with conventional computing architectures.

Various attempts have been made to implement recurrent LSTM networks using memristors. Figure 6D shows LSTM networks in memristor CBAs. A Ta/HfO$_2$ ionic migration memristor exhibits a pinched I-V hysteresis loop and distinguishable conductance states. This device is integrated into a 12,8 × 64 1T-1R memristor CBA and implements LSTM networks, as shown in Figure 6E (Li et al., 2019). The width profiles of human silhouettes extracted from a video were entered into the LSTM algorithms to perform learning. The cross-entropy loss decreased during the update processes within the algorithm, and it was confirmed that the accuracy of identifying humans based on their gaits reached 79.1%, after 50 training epochs. In addition, the replication of neurobiological synaptic characteristics, including short-term plasticity and long-term plasticity, in studies has been reported (Chang et al., 2011; Kumar et al., 2017; Nayak et al., 2012).
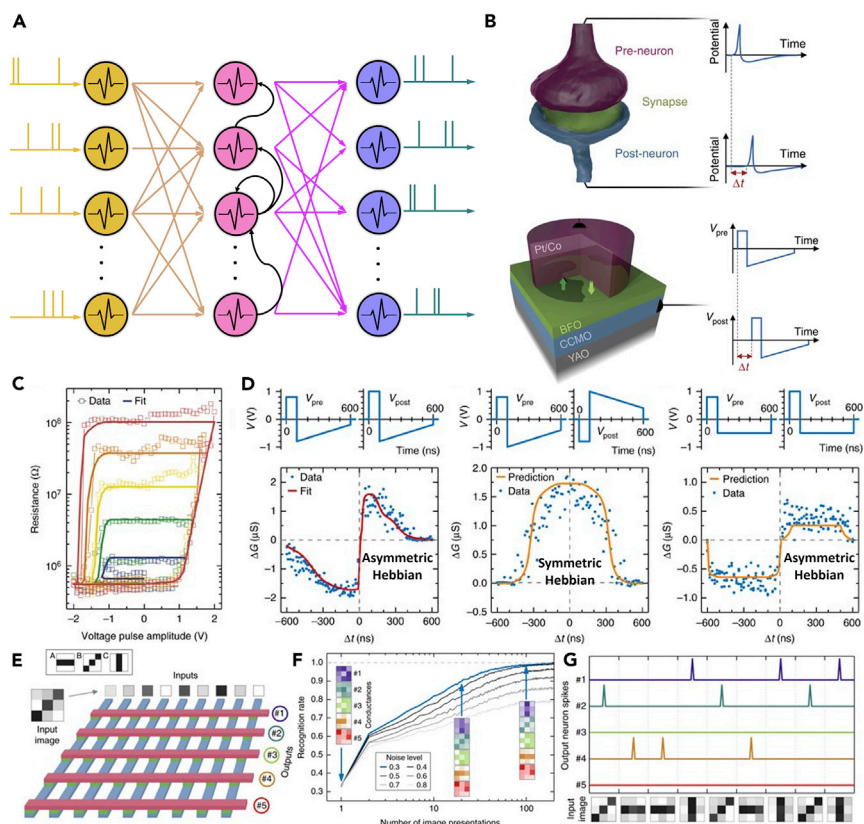
Spiking neural networks with memristors

SNNs are completely different from the other neural networks described above. Unlike other neural networks, the SNN is reported to be the most human-brain-like algorithm. In biological neural systems, neurons are only activated when they detect an input signal. For the rest, neurons are in a resting state, thereby saving enormous energy. This type of computation is called event-driven processing. The SNN emulates spiking neurons using the leaky-integrated-and-fire (LIF) model. The LIF model mimics the action potential of biological neurons, which are fired when the threshold voltage is reached, and the resting potential is recovered. Although other neural network algorithms process the information using floating-point numbers, the SNN encodes the data into the pulse timing and frequency between the pre- and post-synaptic spikes. These characteristics are more similar to actual biological nervous systems and are more energy efficient than other neural networks. Therefore, the SNN is considered to be the next neural network algorithm for ANNs (Roy et al., 2019). Figure 7A illustrates the schematic diagram of a multi-layer SNN.

The fundamentally requisite synaptic characteristics of SNN are the diffusion dynamics of biological neural systems since SNN transmits information through the time difference between the pre- and post-synaptic neuron's spikes. Repeated action potential leads to increase Ca^{2+} in pre-synaptic neurons and induce to release of more neurotransmitters. They cause higher excitatory post-synaptic potential and current (EPSP or EPSC). Accordingly, the nervous systems transmit the signal and memorize through the diffusion of ions. The representative phenomena are paired-pulse facilitation (regarded as short-term plasticity since the mechanism is based on the depletion of neurotransmitter vesicles) and spike-timing-dependent plasticity (STDP).

Especially, the STDP is long-term synaptic plasticity that updates synaptic weight, according to the relative timing between the pre- and post-neuron spiking (Miller et al., 2000). This originated from Hebb's learning rule. In 1950, Hebb proposed the basic physiological learning principles based on synaptic plasticity, known as Hebb's learning rule (Hebb, 2005). The strength of the synapse connectivity can be adjusted based on the signal between pre- and post-neuron firing. This theory is often summarized as "neurons that fire together wire together."

Figure 7. Spiking neural networks with memristors

(A) Schematic of multi-layer SNNs.

(B) Schematics of biological synapse and ferroelectric memristor where pre- and post-neurons are connected.

(C) Resistive voltage multiple hysteresis loops of a ferroelectric memristor show dependence on maximum voltage pulse amplitude.

(D and E) (D) Spike-time-dependent plasticity learning curves for different programmed voltage pulses (E) 9 × 5 CBA to simulate SNNs.

(F) Recognition accuracy rate by unsupervised learning iteration.

(G) Output spiking neuron inference of the network after successful training.

Reproduced with permission (Boyn et al., 2017) Copyright 2017, Springer Nature.

The realization of the STDP has been demonstrated in several memristor studies. Boyn et al. fabricated ferroelectric memristors to implement SNNs (Boyn et al., 2017). Figure 7B shows a sketch of spike-based pre- and post-neurons connected by an artificial synapse. A ferroelectric layer of $BiFeO_3$ is sandwiched between a Pt/Co top electrode, and a $CaCeMnO_3$ bottom electrode stands on $YAlO_3$. As seen in Figure 7C, the ferroelectric memristor shows multiple hysteresis loops in an R-V characteristic, indicating a dependence of the maximum pulse amplitude. Based on this memristive characteristic, this device exhibited STDP learning of different shapes. Figure 7D shows that the programmed pre- and post-spiking voltage pulses are transmitted to each electrode, and the conductance variation depends on the relative spiking timing. These data are also conformed to the asymmetric Hebb's rule and the symmetric Hebb's rule. Boyn et al. simulated the SNN, which has nine input neurons connected to five output neurons based on the CBA. Figure 7E illustrates the CBA of the ferroelectric memristors. The input images are three types of noisy image patterns: horizontal, diagonal, and vertical 3 × 3-pixel images. As seen in Figure 7F, the recognition accuracy reaches approximately 99% for low noise levels and 80% for high noise levels after 100 epochs. After successful learning, the network classifies the input images with high accuracy. Moreover, numerous types of memristors, such as ionic migration (Jo et al., 2010; Kim et al., 2015; Yang et al., 2018c), phase change (Kuzum et al., 2012; Li et al., 2014; Suri et al., 2011), spin (Srinivasan et al., 2016), and ferroelectric tunnel junctions (Chen et al., 2018; Nau, 2017; Yoong et al., 2018), have been reported to implement the STDP for SNN operation. Furthermore, Wang et al. demonstrate the diffusive memristors with Ag-in-oxide thin film to emulate the diffusive dynamics of biological neurons (Wang et al., 2017).
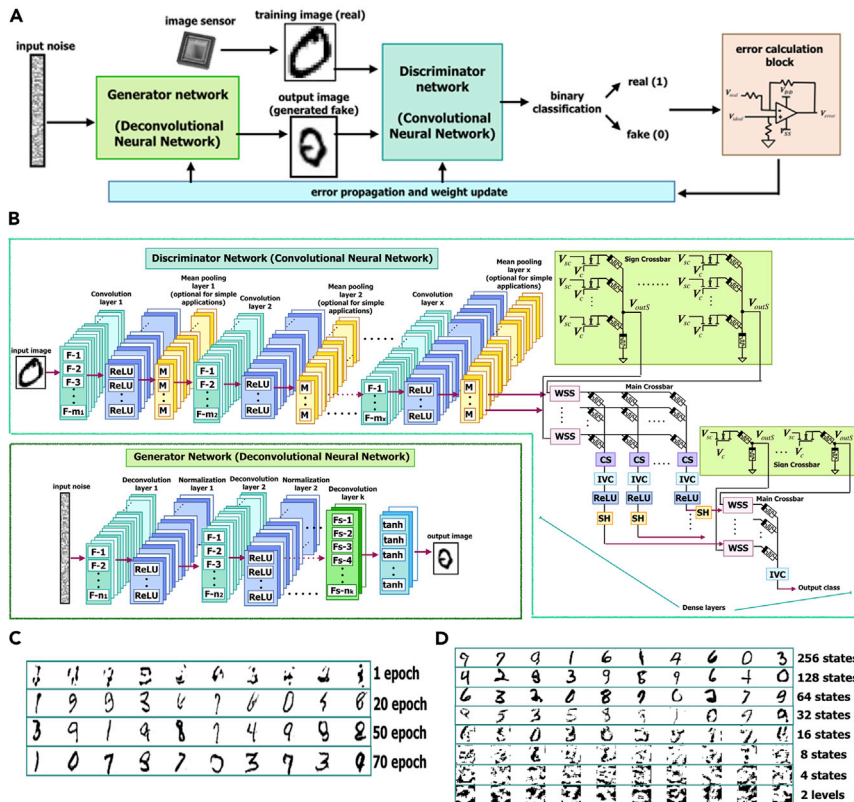
Figure 8. Generative adversarial networks with memristors

(A) Flow chart of analog memristive deep convolution generative adversarial networks.

(B) Discriminator architecture consisting of convolutional neural network and generator architecture consisting of a deconvolutional neural network.

(C) Images generated by the number of training iterations.

(D) Effect of the number of distinguishable conductance levels in memristor unit on generated images.

Reproduced with permission (Krestinskaya et al., 2020) Copyright 2018, Springer Nature.

Generative adversarial networks with memristors

One notable feature of human intelligence is creativity. Human beings can create new things that never existed, based on learned information. They can also relearn their created results and design new ones. In neuromorphic computing systems, GANs can generate new data through unsupervised learning.

The GAN was first proposed by Goodfellow et al. in 2014 (Goodfellow et al., 2014). They proposed two neural networks and pitted them against each other in a game. Figure 8A illustrates a typical deep convolutional GAN architecture with analog memristive systems. The generator network generates fake image data candidates through deconvolutional processing, and the discriminator network distinguishes the candidates from the true data distributions by the CNN, compared to training real images. Through error propagation and weight update in the neural network, the generator network creates data that are closer to the real images. In addition, the GAN translates an image from one into the other (e.g. winter to summer, horse to zebra, painting style of Monet to Van Gogh.) (Zhu et al., 2017).

Although this field has not been extensively studied, significant studies have been conducted to implement this algorithm using a memristor CBA. Krestinskaya et al. proposed an analog memristive deep convolution GAN (AM-DCGAN) architecture with WO_x resistive switching devices in 2020 (Krestinskaya et al., 2020). Figure 8B presents the overall architecture of the AM-DCGAN. They designed a discriminator network using a CNN and a generator network using a DNN. Through this network, distinguishable handwritten number images were generated after 70 epochs of training (shown in Figure 8C). Furthermore, they performed simulations to confirm the effect of the number of distinguishable conductance

levels in the memristor unit on the generated image quality. Figure 8D shows the images that were generated by learning the MNIST handwritten images in the GAN under the condition that 2 to 256 distinct states are stored in one memristor unit. This confirms that the memristor achieves more distinguishable conductance states and the GAN can produce more accurate images. In addition, other memristors, including ionic migration (Chen et al., 2019; Lin et al., 2019) and spin (Dong et al., 2019), and a new kind of GAN (Roohi et al., 2020), have been studied to achieve more creativity in neuromorphic computing systems.

CONCLUSION AND PERSPECTIVES

The massive data produced by numerous edge devices and sensors have supported the striking development of AI. Several research institutes have estimated that the total global data amount will grow to 175 zettabytes by 2025 (Reinsel et al., 2018). This has necessitated higher-performance computing power to process big data. However, it has been predicted that the conventional CMOS-based computing systems will reach their physical and performance limits by 2024, according to the 2018 International Roadmap for Devices and Systems. These facts necessitate a paradigm shift to post-CMOS technology to overcome volatility, multi-state parallel operations, and von Neumann bottlenecks.

Intensive research has demonstrated that memristor-based hardware achieves groundbreaking performance in neuromorphic computing, which is inspired by the human brain. They are considered excellent candidates for alternate conventional computing systems. However, despite their considerable potential, there are still material-, device-, architecture-, and algorithm-level challenges impeding their commercialization.

Compared to the development of AI software, learning-based AI hardware is lagging far behind. Thus far, researchers on memristor materials have found it difficult to manipulate reliable multi-level states, where the memristor unit can be sufficiently distinguished. Therefore, the commercialization of memristors for AI will be accelerated through the development of high-performance and reliable materials, the standardization of synaptic properties, such as conventional CMOS technologies, and road mapping of related technologies. Recently, memristor-CMOS hybrid AI hardware has yielded good performance (Thakur et al., 2018). These hybrid computing systems are considered to be effective for boosting the development of neuromorphic computing systems because CMOS technology has been well developed. Besides developing memristor devices based on a specific material system, hybridizing various memristive materials on a system can be an alternative path to actualizing earlier commercialization.

Furthermore, an architecture that completely controls the sneak current in the CBA structure is required. Large-scale size studies and the existing computing architectures remain hindered. In addition, neuroscience needs to clarify the learning mechanism of the human brain to develop high-performance neural network algorithms.

In summary, significant research has been conducted to address these limitations, and the memristor technology, which is a building block of neuromorphic computing systems, is expected to give rise to new computing systems. In the near future, we hope that computers inspired by the human brain will outperform the human brain.

AUTHOR CONTRIBUTIONS

S.B.K. and H.W.J. provided the outline and conducted writing. S.J.K. conducted the literature review, wrote, and revised the manuscript. All authors discussed the results and contributed to the final manuscript.

REFERENCES

Ambrogio, S., Narayanan, P., Tsai, H., Shelby, R.M., Boybat, I., Di Nolfo, C., Sidler, S., Giordano, M., Bodini, M., Farinha, N.C.P., et al. (2018). Equivalent-accuracy accelerated neural-network training using analogue memory. Nature 558, 60–67.

Atkinson, R.C., and Shiffrin, R.M. (1968). Human memory: a proposed system and its control processes. Psychol. Learn. Motiv. 2, 89–195.

Azevedo, F.A.C., Carvalho, L.R.B., Grinberg, L.T., Farfel, J.M., Ferretti, R.E.L., Leite, R.E.P., Filho, W.J., Lent, R., and Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J. Comp. Neurol. 513, 532–541.

Baibich, M.N., Broto, J.M., Fert, A., Van Dau, F.N., Petroff, F., Eitenne, P., Creuzet, G., Friederich, A., and Chazelas, J. (1988). Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. Phys. Rev. Lett. 61, 2472–2475.

La Barbera, S., Vuillaume, D., and Alibart, F. (2015). Filamentary switching: synaptic plasticity through device volatility. ACS Nano 9, 941–949.

Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Networks 5, 157–166.

Böscke, T.S., Müller, J., Bräuhaus, D., Schröder, U., and Böttger, U. (2011). Ferroelectricity in hafnium oxide thin films. Appl. Phys. Lett. 99, 0–3.

Boyn, S., Grollier, J., Lecerf, G., Xu, B., Locatelli, N., Fusil, S., Girod, S., Carrétéro, C., Garcia, K., Xavier, S., et al. (2017). Learning through ferroelectric domain dynamics in solid-state synapses. Nat. Commun. 13, 878.

Brataas, A., and Hals, K.M.D. (2014). Spin-orbit torques in action. Nat. Nanotechnol. 9, 86–88.

Burr, G.W., Shenoy, R.S., Virwani, K., Narayanan, P., Padilla, A., Kurdi, B., and Hwang, H. (2014). Access devices for 3D crosspoint memory. J. Vac. Sci. Technol. B 32, 040802.

Burr, G.W., Shelby, R.M., Sebastian, A., Kim, S., Kim, S., Sidler, S., Virwani, K., Ishii, M., Narayanan, P., Fumarola, A., et al. (2017). Neuromorphic computing using non-volatile memory. Adv. Phys. X 2, 89–124.

Cai, F., Correll, J.M., Lee, S.H., Lim, Y., Bothra, V., Zhang, Z., Flynn, M.P., and Lu, W.D. (2019). A fully integrated reprogrammable memristor–CMOS system for efficient multiply–accumulate operations. Nat. Electron. 2, 290–299.

Caterini, A.L., and Chang, D.E. (2018). Deep Neural Networks in a Mathematical Framework (Springer).

Chang, T., Jo, S.H., and Lu, W. (2011). Short-term memory to long-term memory transition in a nanoscale memristor. ACS Nano 5, 7669–7676.

Chanthbouala, A., Crassous, A., Garcia, V., Bouzehouane, K., Fusil, S., Moya, X., Allibe, J., Dlubak, B., Grollier, J., Xavier, S., et al. (2012). Solid-state memories based on ferroelectric tunnel junctions. Nat. Nanotechnol. 7, 101–104.

Chen, F., Song, L., and Li, H.H. (2019). Efficient process-in-memory architecture design for unsupervised GAN-based deep learning using ReRAM. Proc. ACM Gt. Lakes Symp. 423–428.

Chen, L., Wang, T.Y., Dai, Y.W., Cha, M.Y., Zhu, H., Sun, Q.Q., Ding, S.J., Zhou, P., Chua, L., and Zhang, D.W. (2018). Ultra-low power Hf0.5Zr0.5O2 based ferroelectric tunnel junction synapses for hardware neural network applications. Nanoscale 10, 15826–15833.

Cho, B., Yun, J.M., Song, S., Ji, Y., Kim, D.Y., and Lee, T. (2011). Direct observation of Ag filamentary paths in organic resistive memory devices. Adv. Funct. Mater. 21, 3976–3981.

Choi, B.J., Yang, J.J., Zhang, M.X., Norris, K.J., Ohlberg, D.A.A., Kobayashi, N.P., Medeiros-Ribeiro, G., and Williams, R.S. (2012). Nitride memristors. Appl. Phys. A 109, 1–4.

Choi, B.J., Torrezan, A.C., Norris, K.J., Miao, F., Strachan, J.P., Zhang, M.X., Ohlberg, D.A.A., Kobayashi, N.P., Yang, J.J., and Williams, R.S. (2013). Electrical performance and scalability of Pt dispersed SiO2 nanometallic resistance switch. Nano Lett. 13, 3213–3217.

Choi, J., Park, S., Lee, J., Hong, K., Kim, D.H., Moon, C.W., Park, G. Do, Suh, J., Hwang, J., Kim, S.Y., et al. (2016). Organolead halide perovskites for low operating voltage multilevel resistive switching. Adv. Mater. 28, 6562–6567.

Choi, Y., Lee, C., Kim, M., Song, Y., Hwang, H., and Lee, D. (2019). Structural engineering of Li-based electronic synapse for high reliability. IEEE Electron. Device Lett. 40, 1992–1995.

Chua, L. (1971). Memristor. The missinf. IEEE Trans. Circuit Theor. 18, 507–519.

Chua, L. (2014). If it's pinched it's a memristor. Semicond. Sci. Technol. 29, 104001.

Dennard, R., Gaensslen, F., Yu, W.-N., Rideout, L., Bassous, E., and Le Blanc, A. (1974). Design of ion-implanted small MOSFET ' S dimensions with very. IEEE J. Solid State Circuits 9, 257–268.

Dieny, B., Speriosu, V.S., Parkin, S.S.P., Gurney, B.A., Wilhoit, D.R., and Mauri, D. (1991). Giant magnetoresistive in soft ferromagnetic multilayers. Phys. Rev. B 43, 1297–1300.

Dong, Z., Fang, Y., Huang, L., Li, J., and Qi, D. (2019). A compact memristor-based GAN architecture with a case study on single image super-resolution. In Procedings of 31st Chinese Control and Decision Conference CCDC 2019, pp. 3069–3074.

Drachman, D.A. (2005). Do we have brain to spare? Neurology 64, 2004–2005.

Dünkel, S., Trentzsch, M., Richter, R., Moll, P., Fuchs, C., Gehring, O., Majer, M., Wittek, S., Müller, B., Melde, T., et al. (2018). A FeFET based super-low-power ultra-fast embedded NVM technology for 22nm FDSOI and beyond. Tech. Dig. Int. Electron. Devices Meet. 19, 7.1–19.7.4.

Fan, D., Shim, Y., Raghunathan, A., and Roy, K. (2015). STT-SNN: a spin-transfer-torque based soft-Limiting non-linear neuron for low-power artificial neural networks. IEEE Trans. Nanotechnol. 14, 1013–1023.

Fuller, E.J., Gabaly, F. El, Léonard, F., Agarwal, S., Plimpton, S.J., Jacobs-Gedrim, R.B., James, C.D., Marinella, M.J., and Talin, A.A. (2017). Li-ion synaptic transistor for low power analog computing. Adv. Mater. 29, 1–8.

Garcia, V., Fusil, S., Bouzehouane, K., Enouz-Vedrenne, S., Mathur, N.D., Barthélémy, A., and Bibes, M. (2009). Giant tunnel electroresistance for non-destructive readout of ferroelectric states. Nature 460, 81–84.

Ge, C., Liu, C.X., Zhou, Q.L., Zhang, Q.H., Du, J.Y., Li, J.K., Wang, C., Gu, L., Yang, G.Z., and Jin, K.J. (2019). A ferrite synaptic transistor with topotactic transformation. Adv. Mater. 31, 1–9.

Gerasimov, J.Y., Gabrielsson, R., Forchheimer, R., Stavrinidou, E., Simon, D.T., Berggren, M., and Fabiano, S. (2019). An evolvable organic electrochemical transistor for neuromorphic applications. Adv. Sci. 6, 1–8.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. Adv. Neural Inf. Process. Syst. 2672–2680.

Grollier, J., Querlioz, D., and Stiles, M.D. (2016). Spintronic nanodevices for bioinspired computing. Proc. IEEE 104, 2024–2039.

Grollier, J., Querlioz, D., Camsari, K.Y., Everschor-Sitte, K., Fukami, S., and Stiles, M.D. (2020). Neuromorphic spintronics. Nat. Electron. 3, 1–11.

Hady, F.T., Foong, A., Veal, B., and Williams, D. (2017). Platform storage performance with 3D XPoint technology. Proc. IEEE 105, 1822–1833.

Han, J.S., Le, Q. Van, Choi, J., Kim, H., Kim, S.G., Hong, K., Moon, C.W., Kim, T.L., Kim, S.Y., and Jang, H.W. (2019). Lead-free all-inorganic cesium tin iodide perovskite for filamentary and interface-type resistive switching toward environment-friendly and temperature-tolerant nonvolatile memories. ACS Appl. Mater. Inter. 11, 8155–8163.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification (IEEE international conference on computer vision(ICCV)), pp. 1026–1034, arXiv:1502.01852v1.

Hebb, D.O. (2005). The Organization of Behavior: A Neuropsychological Theory (Psychology Press).

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. Neural Comput. 9, 1735–1780.

Hsu, C., Wang, I., Lo, C., Chiang, M., and Jang, W. (2013). Self-rectifying bipolar TaOx/TiO2 RRAM with superior endurance over 1012 cycles for 3D high-density storage-class memory. In 2013 Symposium on VLSI Technology (IEEE), pp. T166–T167.

Huang, Y., Kang, W., Zhang, X., Zhou, Y., and Zhao, W. (2017). Magnetic skyrmion-based synaptic devices. Nanotechnology 28, 8.

Im, I.H., Kim, S.J., and Jang, H.W. (2020). Memristive devices for new computing paradigms. Adv. Intell. Syst. *2*, 2000105.

Ioannou, P.S., Kyriakides, E., Schneegans, O., and Giapintzakis, J. (2020). Evidence of biorealistic synaptic behavior in diffusive Li-based two-terminal resistive switching devices. Sci. Rep. *10*, 1–10.

Jerry, M., Chen, P.Y., Zhang, J., Sharma, P., Ni, K., Yu, S., and Datta, S. (2017). Ferroelectric FET analog synapse for acceleration of deep neural network training. IEEE International Electron Devices Meeting (IEDM), pp. 6.2.1–6.2.4.

Jiang, H., Han, L., Lin, P., Wang, Z., Jang, M.H., Wu, Q., Barnell, M., Yang, J.J., Xin, H.L., and Xia, Q. (2016). Sub-10 nm Ta channel responsible for superior performance of a HfO2 memristor. Sci. Rep. *6*, 28525.

Jo, S.H., Chang, T., Ebong, I., Bhadviya, B.B., Mazumder, P., and Lu, W. (2010). Nanoscale memristor device as synapse in neuromorphic systems. Nano Lett. *10*, 1297–1301.

Jude Hemanth, D., Vijila, C.K.S., Selvakumar, A.I., and Anitha, J. (2014). Performance improved iteration-free artificial neural networks for abnormal magnetic resonance brain image classification. Neurocomputing *130*, 98–107.

Kim, S., Du, C., Sheridan, P., Ma, W., Choi, S., and Lu, W.D. (2015). Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity. Nano Lett. *15*, 2203–2211.

Kim, S., Kim, H., Hwang, S., Kim, M.H., Chang, Y.F., and Park, B.G. (2017). Analog synaptic behavior of a silicon nitride memristor. ACS Appl. Mater. Inter. *9*, 40420–40427.

Kim, S.G., Van Le, Q., Han, J.S., Kim, H., Choi, M.J., Lee, S.A., Kim, T.L., Kim, S.B., Kim, S.Y., and Jang, H.W. (2019). Dual-phase All-inorganic cesium halide perovskites for conducting-bridge memory-based artificial synapses. Adv. Funct. Mater. *29*, 1–11.

Krestinskaya, O., Choubey, B., and James, A.P. (2020). Memristive GAN in analog. Sci. Rep. *10*, 5838.

Kumar, S., Strachan, J.P., and Williams, R.S. (2017). Chaotic dynamics in nanoscale NbO 2 Mott memristors for analogue computing. Nature *548*, 318–321.

Kuzum, D., Jeyasingh, R.G.D., Lee, B., and Wong, H.S.P. (2012). Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. Nano Lett. *12*, 2179–2186.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. Proc. IEEE *86*, 2278–2323.

Li, C., Hu, M., Li, Y., Jiang, H., Ge, N., Montgomery, E., Zhang, J., Song, W., Dávila, N., Graves, C.E., et al. (2018). Analogue signal and image processing with large memristor crossbars. Nat. Electron. *1*, 52–59.

Li, C., Wang, Z., Rao, M., Belkin, D., Song, W., Jiang, H., Yan, P., Li, Y., Lin, P., Hu, M., et al. (2019).

Long short-term memory networks in memristor crossbar arrays. Nat. Mach. Intell. *1*, 49–57.

Li, J., Ge, C., Du, J., Wang, C., Yang, G., and Jin, K. (2020). Reproducible ultrathin ferroelectric domain switching for high-performance neuromorphic computing. Adv. Mater. *32*, 1–9.

Li, S., Zeng, F., Chen, C., Liu, H., Tang, G., Gao, S., Song, C., Lin, Y., Pan, F., and Guo, D. (2013). Synaptic plasticity and learning behaviours mimicked through Ag interface movement in an Ag/conducting polymer/Ta memristive system. J. Mater. Chem. C *1*, 5292–5298.

Li, Y., Zhong, Y., Zhang, J., Xu, L., Wang, Q., Sun, H., Tong, H., Cheng, X., and Miao, X. (2014). Activity-dependent synaptic plasticity of a chalcogenide electronic synapse for neuromorphic systems. Sci. Rep. *4*, 1–7.

Lin, P., Li, C., Wang, Z., Li, Y., Jiang, H., Song, W., Rao, M., Zhuo, Y., Upadhyay, N.K., Barnell, M., et al. (2020). Three-dimensional memristor circuits as complex neural networks. Nat. Electron. *3*, 225–232.

Lin, Y., Wu, H., Gao, B., Yao, P., Wu, W., Zhang, Q., Zhang, X., Li, X., Li, F., Lu, J., et al. (2018). Demonstration of generative adversarial network by intrinsic random noises of analog rram devices (IEEE International Electron Devices Meeting (IEDM)), pp. 3.4. 1–3.4. 4.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., and Sánchez, C.I. (2017). A survey on deep learning in medical image analysis. Med. Image Anal. *42*, 60–88.

Liu, L., Pai, C.F., Li, Y., Tseng, H.W., Ralph, D.C., and Buhrman, R.A. (2012). Spin-torque switching with the giant spin hall effect of tantalum. Science *336*, 555–558.

Locatelli, N., Cros, V., and Grollier, J. (2014). Spin-torque building blocks. Nat. Mater. *13*, 11–20.

Luo, Z.D., Peters, J.J.P., Sanchez, A.M., and Alexe, M. (2019). Flexible memristors based on single-crystalline ferroelectric tunnel junctions. ACS Appl. Mater. Inter. *11*, 23313–23319.

Mai, V.H., Moradpour, A., Senzier, P.A., Pasquier, C., Wang, K., Rozenberg, M.J., Giapintzakis, J., Mihailescu, C.N., Orfanidou, C.M., Svoukis, E., et al. (2015). Memristive and neuromorphic behavior in a Li x CoO 2 nanobattery. Sci. Rep. *5*, 7761.

Majumdar, S., Tan, H., Qin, Q.H., and van Dijken, S. (2019). Energy-efficient organic ferroelectric tunnel junction memristors for neuromorphic computing. Adv. Electron. Mater. *5*, 1–10.

McCulloch, W.S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. *5*, 115–133.

Mehonic, A., and Kenyon, A.J. (2016). Emulating the electrical activity of the neuron using a silicon oxide RRAM cell. Front. Neurosci. *10*, 57.

Miller, K.D., Abbott, L.F., and Song, S. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat. Neurosci. *3*, 919–926.

Mohamed, A.R., Dahl, G.E., and Hinton, G. (2012). Acoustic modeling using deep belief networks. IEEE Trans. Audio, Speech Lang. Process. *20*, 14–22.

Moore, G.M. (1965). Moore's law ,electronics. Electronics *38*, 114.

Moradpour, A., Schneegans, O., Franger, S., Revcolevschi, A., Salot, R., Auban-Senzier, P., Pasquier, C., Svoukis, E., Giapintzakis, J., Dragos, O., et al. (2011). Resistive switching phenomena in Li xCoO 2 thin films. Adv. Mater. *23*, 4141–4145.

Mueller, S., Mueller, J., Singh, A., Riedel, S., Sundqvist, J., Schroeder, U., and Mikolajick, T. (2012). Incipient ferroelectricity in Al-doped HfO 2 thin films. Adv. Funct. Mater. *22*, 2412–2417.

Mulaosmanovic, H., Chicca, E., Bertele, M., Mikolajick, T., and Slesazeck, S. (2018). Mimicking biological neurons with a nanoscale ferroelectric transistor. Nanoscale *10*, 21755–21763.

Müller, J., Schröder, U., Böscke, T.S., Müller, I., Böttger, U., Wilde, L., Sundqvist, J., Lemberger, M., Kücher, P., Mikolajick, T., et al. (2011). Ferroelectricity in yttrium-doped hafnium oxide. J. Appl. Phys. *110*, 1–6.

Nau, J.Y. (2017). Miscellanées synaptiques, néo-zélandaises et électroniques. Rev. Med. Suisse *13*, 878.

Nayak, A., Ohno, T., Tsuruoka, T., Terabe, K., Hasegawa, T., Gimzewski, J.K., and Aono, M. (2012). Controlling the synaptic plasticity of a Cu 2 S gap-type atomic switch. Adv. Funct. Mater. *22*, 3606–3613.

Oh, S., Hwang, H., and Yoo, I.K. (2019). Ferroelectric materials for neuromorphic computing. APL Mater. *7*, 091109.

Ovshinsky, S.R. (1968). Reversible electrical switching phenomena in disordered structures. Phys. Rev. Lett. *21*, 1450–1453.

Pirovano, A., Lacaita, A.L., Benvenuti, A., Pellizzer, F., and Bez, R. (2004). Electronic switching in phase-change memories. IEEE Trans. Electron. Devices *51*, 452–459.

Reinsel, D., Gantz, J., and Rydning, J. (2018). The Digitization of the World - from Edge to Core (IDC).

Roohi, A., Sheikhfaal, S., Angizi, S., Fan, D., and Demara, R.F. (2020). ApGAN: Approximate GAN for robust low energy learning from imprecise components. IEEE Trans. Comput. *69*, 349–360.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. *65*, 386–408.

Roy, K., Jaiswal, A., and Panda, P. (2019). Towards spike-based machine intelligence with neuromorphic computing. Nature *575*, 607–617.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv, 1609.04747.

Sasaki, T., Ukyo, Y., and Novák, P. (2013). Memory effect in a lithium-ion battery. Nat. Mater. *12*, 569–575.

Scott, J.F., and Paz De Araujo, C.A. (1989). Ferroelectric memories. Science 246, 1400–1405.

Sharad, M., Fan, D., and Roy, K. (2013). Spin-neurons: a possible path to energy-efficient neuromorphic computers. J. Appl. Phys. 114, 234906.

Sharbati, M.T., Du, Y., Torres, J., Ardolino, N.D., Yun, M., and Xiong, F. (2018). Low-power, electrochemically tunable graphene synapses for neuromorphic computing. Adv. Mater. 30, 1–6.

Shi, J., Ha, S.D., Zhou, Y., Schoofs, F., and Ramanathan, S. (2013). A correlated nickelate synaptic transistor. Nat. Commun. 4, 1–9.

Siegrist, T., Jost, P., Volker, H., Woda, M., Merkelbach, P., Schlockermann, C., and Wuttig, M. (2011). Disorder-induced localization in crystalline phase-change materials. Nat. Mater. 10, 202–208.

Slonczewski, J.C. (1996). Current-driven excitation of magnetic multilayers. J. Magn. Magn. Mater. 159, L1–L7.

Song, K.M., Jeong, J.S., Pan, B., Zhang, X., Xia, J., Cha, S., Park, T.E., Kim, K., Finizio, S., Raabe, J., et al. (2020). Skyrmion-based artificial synapses for neuromorphic computing. Nat. Electron. 3, 148–155.

Srinivasan, G., Sengupta, A., and Roy, K. (2016). Magnetic tunnel junction based long-term short-term stochastic synapse for a spiking neural network with on-chip STDP learning. Sci. Rep. 6, 29545.

Stengel, M., Vanderbilt, D., and Spaldin, N.A. (2009). Enhancement of ferroelectricity at metal-oxide interfaces. Nat. Mater. 8, 392–397.

Strukov, D.B., Snider, G.S., Stewart, D.R., and Williams, R.S. (2008). The missing memristor found. Nature 453, 80–83.

Sun, H., Liu, Q., Li, C., Long, S., Lv, H., Bi, C., Huo, Z., Li, L., and Liu, M. (2014). Direct observation of conversion between threshold switching and memory switching induced by conductive filament morphology. Adv. Funct. Mater. 24, 5679–5686.

Suri, M., Bichler, O., Querlioz, D., Cueto, O., Perniola, L., Sousa, V., Vuillaume, D., Gamrat, C., and DeSalvo, B. (2011). Phase change memory as synapse for ultra-dense neuromorphic systems: application to complex visual pattern extraction. In 2011 International Electron Devices Meeting (IEEE), pp. 4.4.1–4.4.4.

Suri, M., Bichler, O., Querlioz, D., Traoré, B., Cueto, O., Perniola, L., Sousa, V., Vuillaume, D., Gamrat, C., Desalvo, B., et al. (2012). Physical aspects of low power synapses based on phase change memory devices. J. Appl. Phys. 112, 054904.

Sze, V., Chen, Y.-H., Yang, T.-J., and Emer, J.S. (2017). MIT tutorial. Proc. IEEE 105, 2295–2329.

Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. (2014). DeepFace: closing the gap to human-level performance in face verification. In 2014 IEEE Conference on Computer Vision and Pattern Recognition (IEEE), pp. 1701–1708.

Tang, J., Yuan, F., Shen, X., Wang, Z., Rao, M., He, Y., Sun, Y., Li, X., Zhang, W., Li, Y., et al. (2019). Bridging biological and artificial neural networks with emerging neuromorphic devices: fundamentals, progress, and challenges. Adv. Mater. 31, 092001.

Thakur, C.S., Molin, J.L., Cauwenberghs, G., Indiveri, G., Kumar, K., Qiao, N., Schemmel, J., Wang, R., Chicca, E., Olson Hasler, J., et al. (2018). Large-scale neuromorphic spiking array processors: a quest to mimic the brain. Front. Neurosci. 12, 1–37.

Tominaga, J., Kikukawa, T., Takahashi, M., and Phillips, R.T. (1997). Structure of the optical phase change memory alloy, Ag-V-In-Sb-Te, determined by optical spectroscopy and electron diffraction. J. Appl. Phys. 82, 3214–3218.

Tsymbal, E.Y., and Kohlstedt, H. (2006). Tunneling across a ferroelectric. Science 313, 181–183.

Tuma, T., Pantazi, A., Le Gallo, M., Sebastian, A., and Eleftheriou, E. (2016). Stochastic phase-change neurons. Nat. Nanotechnol. 11, 693–699.

Turing, A.M. (1936). [1936][Journal of math, allen turing] on computable numbers, with an application to the Entscheidungsproblem.pdf. J. Math. 58, 345–363.

Vourkas, I., and Sirakoulis, G.C. (2015). Memristor-Based Nanoelectronic Computing Circuits and Architectures (Springer).

Wang, J.J., Hu, S.G., Zhan, X.T., Yu, Q., Liu, Z., Chen, T.P., Yin, Y., Hosaka, S., and Liu, Y. (2018a). Handwritten-digit recognition by hybrid convolutional neural network based on HfO2 memristive spiking-neuron. Sci. Rep. 8, 1–5.

Wang, M., Cai, W., Cao, K., Zhou, J., Wrona, J., Peng, S., Yang, H., Wei, J., Kang, W., Zhang, Y., et al. (2018b). Current-induced magnetization switching in atom-thick tungsten engineered perpendicular magnetic tunnel junctions with large tunnel magnetoresistance. Nat. Commun. 9, 671.

Wang, T.Y., Meng, J.L., He, Z.Y., Chen, L., Zhu, H., Sun, Q.Q., Ding, S.J., Zhou, P., and Zhang, D.W. (2019). Fully transparent, flexible and waterproof synapses with pattern recognition in organic environments. Nanoscale Horizons 4, 1293–1301.

Wang, Z., Yin, M., Zhang, T., Cai, Y., Wang, Y., Yang, Y., and Huang, R. (2016). Engineering incremental resistive switching in TaO: X based memristors for brain-inspired computing. Nanoscale 8, 14015–14022.

Wang, Z., Joshi, S., Savel'ev, S.E., Jiang, H., Midya, R., Lin, P., Hu, M., Ge, N., Strachan, J.P., Li, Z., et al. (2017). Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. Nat. Mater. 16, 101–108.

Wang, Z., Wu, H., Burr, G.W., Hwang, C.S., Wang, K.L., Xia, Q., and Yang, J.J. (2020). Resistive switching materials for information processing. Nat. Rev. Mater. 5, 173–195.

Wei, Z., Kanzawa, Y., Arita, K., Katoh, Y., Kawai, K., Muraoka, S., Mitani, S., Fujii, S., Katayama, K., Iijima, M., et al. (2008). Highly reliable TaO x ReRAM and direct evidence of redox reaction mechanism. In 2008 IEEE International Electron Devices Meeting (IEEE), pp. 1–4.

Woo, J., Lee, D., Koo, Y., and Hwang, H. (2017). Dual functionality of threshold and multilevel resistive switching characteristics in nanoscale HfO2-based RRAM devices for artificial neuron and synapse elements. Microelectron. Eng. 182, 42–45.

Xia, Q., and Yang, J.J. (2019). Memristive crossbar arrays for brain-inspired computing. Nat. Mater. 18, 309–323.

Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M.L., Stolcke, A., Yu, D., and Zweig, G. (2017). Toward human parity in conversational speech recognition. IEEE/ACM Trans. Audio Speech Lang. Process. 25, 2410–2423.

Xu, X.Y., Yin, Z.Y., Xu, C.X., Dai, J., and Hu, J.G. (2014). Resistive switching memories in MoS2 nanosphere assemblies. Appl. Phys. Lett. 104, 033504.

Yamada, E., Ohno, K., and Nishiuchi, N.A. (1991). Rapid-phase transitions of GeTe-Sb2Te3 pseudobinary amorphous thin films for an optical disk memory. J. Appl. Phys. 69, 2849.

Yang, C. Sen, Shang, D.S., Liu, N., Shi, G., Shen, X., Yu, R.C., Li, Y.Q., and Sun, Y. (2017). A synaptic transistor based on quasi-2D molybdenum oxide. Adv. Mater. 29, 1–10.

Yang, C. Sen, Shang, D.S., Liu, N., Fuller, E.J., Agrawal, S., Talin, A.A., Li, Y.Q., Shen, B.G., and Sun, Y. (2018a). All-solid-state synaptic transistor with ultralow conductance for neuromorphic computing. Adv. Funct. Mater. 28, 1–10.

Yang, J.J., Pickett, M.D., Li, X., Ohlberg, D.A.A., Stewart, D.R., and Williams, R.S. (2008). Memristive switching mechanism for metal/oxide/metal nanodevices. Nat. Nanotechnol. 3, 429–433.

Yang, J.T., Ge, C., Du, J.Y., Huang, H.Y., He, M., Wang, C., Lu, H. Bin, Yang, G.Z., and Jin, K.J. (2018b). Artificial synapses emulated by an electrolyte-gated tungsten-oxide transistor. Adv. Mater. 30, 1–10.

Yang, R., Huang, H.M., Hong, Q.H., Yin, X.B., Tan, Z.H., Shi, T., Zhou, Y.X., Miao, X.S., Wang, X.P., Mi, S.B., et al. (2018c). Synaptic suppression triplet-STDP learning rule realized in second-order memristors. Adv. Funct. Mater. 28, 1–10.

Yang, Y., Gao, P., Gaba, S., Chang, T., Pan, X., and Lu, W. (2012). Observation of conducting filament growth in nanoscale resistive memories. Nat. Commun. 3, 732.

Yao, P., Wu, H., Gao, B., Tang, J., Zhang, Q., Zhang, W., Yang, J.J., and Qian, H. (2020). Fully hardware-implemented memristor convolutional neural network. Nature 577, 641–646.

Yoon, J.H., Song, S.J., Yoo, I.H., Seok, J.Y., Yoon, K.J., Kwon, D.E., Park, T.H., and Hwang, C.S. (2014). Highly uniform, electroforming-free, and self-rectifying resistive memory in the Pt/Ta2O5/HfO2-x/TiN structure. Adv. Funct. Mater. 24, 5086–5095.

Yoon, J.H., Kim, K.M., Song, S.J., Seok, J.Y., Yoon, K.J., Kwon, D.E., Park, T.H., Kwon, Y.J., Shao, X., and Hwang, C.S. (2015). Pt/Ta2O5/HfO2-x/Ti resistive switching memory competing with multilevel NAND flash. Adv. Mater. 27, 3811–3816.

CellPress
OPEN ACCESS

Yoong, H.Y., Wu, H., Zhao, J., Wang, H., Guo, R., Xiao, J., Zhang, B., Yang, P., Pennycook, S.J., Deng, N., et al. (2018). Epitaxial ferroelectric Hf0.5Zr0.5O2 thin films and their implementations in memristors for brain-inspired computing. Adv. Funct. Mater. *28*, 1–10.

Young, T., Hazarika, D., Poria, S., and Cambria, E. (2018). Recent trends in deep learning based natural language processing [Review Article]. IEEE Comput. Intell. Mag. *13*, 55–75.

Železný, J., Wadley, P., Olejník, K., Hoffmann, A., and Ohno, H. (2018). Spin transport and spin torque in antiferromagnetic devices. Nat. Phys. *14*, 220–228.

Zhang, S. (2000). Spin Hall effect in the presence of spin diffusion. Phys. Rev. Lett. *85*, 393–396.

Zhang, P., Gao, C., Xu, B., Qi, L., Jiang, C., Gao, M., and Xue, D. (2016). Structural phase transition effect on resistive switching behavior of MoS2-polyvinylpyrrolidone nanocomposites films for flexible memory devices. Small *12*, 2077–2084.

Zhu, J., Yang, Y., Jia, R., Liang, Z., Zhu, W., Rehman, Z.U., Bao, L., Zhang, X., Cai, Y., Song, L., et al. (2018). Ion gated synaptic transistors based on 2D van der Waals crystals with tunable diffusive dynamics. Adv. Mater. *30*, 1–11.

Zhu, J.Y., Park, T., Isola, P., and Efros, A.A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Procedings of IEEE International Conference on Computer Vision. 2017-october, pp. 2242–2251.

Zidan, M.A., Strachan, J.P., and Lu, W.D. (2018). The future of electronics based on memristive systems. Nat. Electron. *1*, 22–29.