Engineering MoSe$_2$/MoS$_2$ heterojunction traps in 2D transistors for multilevel memory, multiscale display, and synaptic functions

Yeonsu Jeong$^{1,3}$, Han Joo Lee$^{1,3}$, Junkyu Park$^2$, Sol Lee$^1$, Hye-Jin Jin$^1$, Sam Park$^1$, Hyunmin Cho$^1$, Sungjae Hong$^1$, Taewook Kim$^1$, Kwanpyo Kim$^1$, Shinhyun Choi$^2$ and Seongil Im$^{1,3*}$

We study a low voltage short pulse operating multilevel memory based on van der Waals heterostack (HS) n-MoSe$_2$/n-MoS$_2$ channel field-effect transistors (FETs). Our HS memory FET exploited the gate voltage ($V_{GS}$)-induced trapping/de-trapping phenomena for Program/Erase functioning, which was maintained for long retention times owing to the existence of heterojunction energy barrier between MoS$_2$ and MoSe$_2$. More interestingly, trapped electron density was incrementally modulated by the magnitude or cycles of a pulsed $V_{GS}$, enabling the HS device to achieve multilevel long-term memory. For a practical demonstration, five different levels of drain current were visualized with multiscale light emissions after our memory FET was integrated into an organic light-emitting diode pixel circuit. In addition, our device was applied to a synapse-imitating neuromorphic memory in an artificial neural network. We regard our unique HS channel FET to be an interesting and promising electron device undertaking multifunctional operations related to the upcoming fourth industrial revolution era.

npj 2D Materials and Applications (2022) 6:23 ; https://doi.org/10.1038/s41699-022-00295-8

INTRODUCTION

In the fourth industrial revolution, which is deeply related to big data, the internet of things (IoT), and artificial intelligence (AI)$^{1-2}$, Si-based device technology has faced imminent challenges for massive data processing and storage because of its own limit in device integration density. To cope with the limit, multivalued logic, multilevel memory, and synaptic memory appear promising ways to secure more efficient data processing and memory capacity. For instance, oxide-based resistive random access memory (RRAM) of two-terminal structures has been studied to achieve such purposes, particularly for synaptic/neuromorphic memory functions$^{3-5}$. Still, it is under development for practical application due to random stochastic filament formation and cross-talk issues. To circumvent the issues of RRAM, three-terminal memories have also been suggested for practical circuit applications$^{6-8}$. Intriguingly, such three-terminal neuromorphic devices can be realized with two-dimensional (2D) transition metal dichalcogenides (TMDs), which have attracted much attention due to their unique physical properties$^{9-12}$. In fact, there have been quite a few studies for bistable (two level: Program and Erase) memory transistors composed of diverse 2D van der Waals (vdW) materials: TMDs, graphene, and h-BN, among others$^{13-17}$. However, multilevel memories in 2D TMD channel-based field-effect transistor (FET) structures are only a few in report and also mostly require high operating voltages and long pulse widths$^{18-22}$ (See Supplementary Table 1).

In the present study, we have fabricated multilevel long-term memory devices practically operating with 60 µs-short pulses at low 7 V based on vdW heterostack (HS) n-TMD FETs, which were extended to multiscale display and neuromorphic memory applications. In principle, all kinds of semiconducting TMDs can be utilized for vdW HS FETs including p-type TMDs. Here, 2D n-MoSe$_2$/n-MoS$_2$ structure was used as a channel of HS FET. Our HS channel FETs initially operate as a simple bistable nonvolatile memory FET with Program and Erase states (PR/ER current ratio $\sim 10^3$), which are defined under positive and negative gate voltage ($V_{GS}$) pulses, respectively. The HS memory device properly exploits the $V_{GS}$-induced trapping/de-trapping phenomena for such Program/Erase functioning, as the trap density-of-state (DOS) of electrons exists at the interface between 2D-like thin MoS$_2$ and MoSe$_2$. Interestingly, the stacking sequence between MoS$_2$ and MoSe$_2$ is found to be important for memory retention performance, which needs a proper heterojunction barrier. Therefore, n-MoS$_2$ in contact with a dielectric serves as the main channel layer for electron conduction in FET, and n-MoSe$_2$ then contacts the source/drain (S/D) electrode. More usefully, trapped electron density could be modulated by the amount of pulsed $V_{GS}$, enabling the HS device to achieve multilevel long-term memory or multivalued data storage. For the multilevel memory functions, our device could distinctively demonstrate at least seven-level states of drain current ($I_D$) by varying $V_{GS}$. The multilevel memory device also nicely went through a cyclic endurance test that enforced >100 pulse cycles; a cycle includes 7 $V_{GS}$ pulses to match seven $I_D$ levels. The seven different levels of $I_D$ were further visualized with multiscale light emissions when our memory FET was integrated into a green organic light-emitting diode (OLED) pixel circuit, which then appears analogous to a conventional gray-scale active matrix pixel circuit composed of two FETs and one OLED. As a final interest, our device was put to neuromorphic operation for synaptic devices in an artificial neural network (ANN), where long-term multilevel memories are necessary. Based on potentiation-depression (P-D) characteristics of our device under multiple 60-µs-short $V_{GS}$ pulses, the simulation results show $\sim$94% recognition accuracy. We regard our unique HS channel FET to be an interesting and promising electron device undertaking...
Materials and device characterization of stacked channel memory FET

Figure 1a illustrates a 3D schematic structure of a MoSe2/MoS2 stacked channel FET with polystyrene (PS)-brush-coated Al2O3 gate insulator and Au gate (G) electrode that is patterned on a glass substrate. High-k Al2O3 is chosen for gate insulator, preferred to conventional SiO2, in consideration of low voltage operation of our nonvolatile HS multilevel memory FET. To form the 2D vdW heterojunction stack channel, a MoS2 flake is first transferred on a PS-brush/Al2O3 gate insulator in sequence. The MoSe2 flake is then stacked on the MoS2, covering the whole area of the bottom MoS2 so that the Au S/D electrode may contact the top surface of MoSe2. An ultrathin PS-brush was used to make the surface of the gate oxide more hydrophobic, which would significantly reduce the trap densities at the channel/gate insulator interface.

RESULTS AND DISCUSSION

Materials and device characterization of stacked channel memory FET

Figure 1a illustrates a 3D schematic structure of a MoSe2/MoS2 stacked channel FET with polystyrene (PS)-brush-coated Al2O3 gate insulator and Au gate (G) electrode that is patterned on a glass substrate. High-k Al2O3 is chosen for gate insulator, preferred to conventional SiO2, in consideration of low voltage operation of our nonvolatile HS multilevel memory FET. To form the 2D vdW heterojunction stack channel, a MoS2 flake is first transferred on a PS-brush/Al2O3 gate insulator in sequence. The MoSe2 flake is then stacked on the MoS2, covering the whole area of the bottom MoS2 so that the Au S/D electrode may contact the top surface of MoSe2. An ultrathin PS-brush was used to make the surface of the gate oxide more hydrophobic, which would significantly reduce the trap densities at the channel/gate insulator interface.
MoSe$_2$ and the MoSe$_2$/MoS$_2$ stack, showing that the local vibrational peaks in the stack are accurately matched to those of a single flake. The electrical properties and transfer characteristics (drain current versus gate-source voltage: $I_D$/$V_G$) of our stack channel FET is shown in Fig. 1f. Our FET exhibits a high ON/OFF $I_D$ ratio of $\sim 10^3$ and ON $I_D$ of a hundred nA at the drain-source voltage, $V_{DS} = 0.1$ V. Clockwise hysteresis appears when a double ($\pm$) $V_{GS}$ sweep was conducted from $-7$ to $7$ V. (The hysteresis window of our device gradually increased by increasing the $V_{GS}$ sweep. See Supplementary Fig. 2 for further details of this device.) Short Program ($7$ V, $60$ μs)/Erase ($-7$ V, $60$ μs) pulse experiments were also conducted on our device. (See Supplementary Fig. 2 for all other results from $60$ μs, $1$ ms, $10$ ms, $100$ ms, $500$ ms, and $1$ s pulses. No significant difference in ON/OFF $I_D$ ratio was observed between $60$ μs and $1$ s pulses.) After the pulse, our device displays a dramatic change in the channel conductivity, showing two distinct states in Fig. 1f (Program/Erase $I_D$ ratio of $\sim 10^5$, note violet vs. black square). Figure 1g displays the cyclic endurance results of our HS memory FET, which were measured during 400 repeated Program/Erase pulse operations. The Program/Erase $I_D$ ratio appears well-endured at $\sim 10^5$. The memory retention properties of our device were also measured, as shown in Fig. 1h, where a high Program/Erase $I_D$ ratio of $\sim 10^5$ is still observed at $V_{DS} = 0.1$ V (and $V_{GS} = -1$ V), even after $10^4$ s elapses since the time of Erase ($-7$ V, $60$ μs) and Program ($+7$ V, $60$ μs) pulses. When the pulse width increases to $1$ s, Program/Erase ratio in retention increases to $10^6$. Consequently, it was confirmed that the stacking channel FET could work as a simple nonvolatile memory device to distinguish two-level data storage at least. On the one hand, the other FET fabricated with a reversed stacking sequence of MoS$_2$/MoSe$_2$ showed much inferior retention of a small Program/Erase $I_D$ ratio of $\sim 40$ only after $500$ s. We attribute this inferior retention behavior to the possibility that the offset-induced barrier at the initial MoSe$_2$/MoS$_2$ heterojunction vanishes at the reversed MoS$_2$/MoSe$_2$ stack junction. In this case, Erase pulse is well functioning to keep its state, but Program state may not be well maintained after pulse owing to some charge leakage to the S/D electrode. The offset barrier effects are well described in the band diagrams of Supplementary Fig. 3. Hence, it is important for nonvolatile memory performance whether the offset-induced barrier exists or not. Such charge leakage results in the rapid increase of Program $I_D$ (see Supplementary Fig. 3d).

**Operation mechanism of stacked channel memory device**

As for the aforementioned memory functions, we suspect any contribution of possible charge traps at the MoSe$_2$/MoS$_2$ heterojunction interface as the origin of such nonvolatile function, although no conspicuous defect signs were observed from the high-resolution STEM image of the interface (Fig. 1d). This is because single-MoS$_2$-channel FET with the same oxide dielectric (PS-brush/Al$_2$O$_3$) hardly shows a large hysteresis, unlike the case of an HS channel FET. When we fabricated a MoS$_2$ single-channel FET with a PS-brush layer, its transfer curve showed only a little hysteresis ($-0.2$ V), as seen in Supplementary Fig. 4. Thus, more information was needed on the possible origin of the memory function in the present FET device. Figure 2a shows an OM image of another MoSe$_2$/MoS$_2$ stack channel memory FET along with scanning Kelvin probe microscopy (SKPM) mapping for both stacked and single-channel areas. The work function of MoSe$_2$ alone appeared to be $4.52$ eV while that of the same MoSe$_2$ but stacked on MoS$_2$ was slightly smaller to be $4.49$ eV, which means that n-MoS$_2$ in the stack area is influenced by n-MoSe$_2$. This is particularly due to electron charge transfer taking place from MoSe$_2$ to MoS$_2$ because of a Fermi level ($E_F$) difference between the two n-type TMDs. Such extra electron charges would increase $E_F$ as transferred to n-MoS$_2$ and some charges should be trapped near or at the MoSe$_2$/MoS$_2$ interface. Based on our SKPM results and the reported band diagram of individual MoS$_2$ and MoSe$_2$, a band diagram of the MoSe$_2$/MoS$_2$ heterojunction could be constructed as seen in Fig. 2b. The interfacial electron trap DOS might come from a mechanical exfoliation and dry-transfer process of 2D TMD flakes while the nature of deep level point defects inside MoS$_2$ is...
also known as related to sulfur vacancies \(^{31}\). Fig. 2c–e and Fig. 2f–h illustrate 2D schematic structures and corresponding band diagrams of our MoSe\(_2\)/MoS\(_2\) stack channel FET, respectively. At the pristine states in Fig. 2c, f, some of the trap DOS is initially occupied by electrons. Under Program pulse \((V_{GS} = 7 \text{ V}, 60 \mu\text{s})\), interface traps and DOS are completely filled with electrons drawn from the S/D electrode, as shown in Fig. 2d, g. Even after the pulse, the filled DOS was effectively maintained without losing many electrons because of the heterojunction energy barrier, as seen in Fig. 2f. As a result, electron accumulation near the MoS\(_2\) channel/dielectric interface becomes electrostatically uneasy due to the trapped electrons. The higher gate voltage is thus required for the dielectric interface to become electrostatically uneasy due to the trapped electrons. Under Erase pulse \((V_{GS} = -7 \text{ V}, 60 \mu\text{s})\), electrons in the trap sites will become de-trapped as shown in Fig. 2e, h. After the de-trapping of electrons, there is no electrostatic hindrance for electron accumulation; the MoS\(_2\) bottom channel of our memory FET is easily controlled by the bottom gate without the trap-induced electrostatic hindrance. \(V_{TH}\) now shifts toward more negative than that of the pristine state. Consequently, channel conductance of the memory FET can be modulated between low conductance (Program) and high conductance states (Erase) by a \(V_{GS}\) pulse.

**Device performance of MoSe\(_2\)/MoS\(_2\) stack channel multilevel memory FET**

More interesting phenomena than the two-state (PR/ER) memory effects would be the multistep modulation of channel conductance, which is obtained by filling or evacuating the trap DOS gradually. Figure 3a shows the transfer curves as measured under a double sweep of \(V_{GS}\) in many steps (seven steps of \(V_{GS}\) from \(±1 \text{ V}\) to \(±7 \text{ V}\)), where \(V_{TH}\) values are defined as \(V_{GS}\) at \(1 \text{nA}\) of \(I_{D}\).

Relations between the \(V_{TH}\) and \(V_{GS}\) sweep range are plotted in Fig. 3b, wherein the \(V_{TH}\) values in all the transfer curve hysteresis appear a little asymmetric but show their initial value fixed at \(-1.5 \text{ V}\) (as indicated by the dark dashed line). Such asymmetry probably results from the fact that the low-energy part of the whole trap DOS area was initially occupied by electrons even before applying \(V_{GS}\) pulses, and the occupied DOS area was larger than the unoccupied area. Hence, a highly negative \(V_{GS}\) (more negative than \(-5 \text{ V}\)) may completely evacuate the trapped electrons from the DOS site (black line) while a relatively small positive \(V_{GS}\) (larger than \(+1.0 \text{ V}\)) attracts the electrons to fill the trap DOS (red line). It is worth noting the inset band diagram, where the initially filled (pristine)-DOS state is indicated by a red line. Our understanding from Fig. 3a, b is that \(V_{TH}\) changes gradually under incremental \(V_{GS}\) steps, which may indicate that charge trapping/de-trapping could also be incrementally controllable by the magnitude of the \(V_{GS}\) pulse. In fact, such incremental \(V_{TH}\) change by a short \(V_{GS}\) pulse (60 \(\mu\text{s}\) to 1 \(\text{ s}\) width) correspondingly comes along with an \(I_{D}\) change in light of transfer characteristics. Figure 3c directly shows the \(I_{D}\) modulation by the magnitude of 60 \(\mu\text{s}\)-short \(V_{GS}\) pulse in our HS memory FET as measured at \(V_{GS}\) of \(-0.5 \text{ V}\) at \(V_{DS} = 0.1 \text{ V}\) for \(I_{D}\) recording. Many \(I_{D}\) states are achievable by applying a \(V_{GS}\) pulse, and here seven states are demonstrated for example ID7, ID6, ID5, ID4, ID3, ID2, and ID1 for \(-7, 2, 3, 4, 5, 6, \text{ and } 7 \text{ V}\) pulses, respectively. Starting from \(30 \text{nA}\) of the ID7 state, \(I_{D}\) decreased to \(200 \text{ fA}\) of ID1 with a pulse voltage change in the above order. Cyclic endurance measurement in Fig. 3d exhibits the same discrete ID7, ID6, ID5, ID4, ID3, ID2, and ID1 states without much variation for 100 cycles, as shown in Fig. 3c. Among the 100 cycles, we zoom into a few cycles to see their details in Fig. 3e, where under cyclic \(V_{GS}\) pulses all seven \(I_{D}\) states are clearly repeated in order. Even if random arbitrary \(V_{GS}\) pulses are also applied, as shown in the upper plots of Fig. 3f, all seven \(I_{D}\) states keep their own original levels...
according to the random order (although two states are skipped here for properly visible demonstration). Moreover, in the case of Fig. 3f, long-term multilevel over 70 s appears to be kept or remembered for each state after a 60 μs VGS pulse, showing that our HS FET plays very well as a multilevel memory device.

**Multilevel light-emitting of OLED pixel with stack channel memory FET**

The above-mentioned long-term multilevel memory functions could be more visibly demonstrated with a practical application such as gray or multiscale OLED pixel switching, for which our MoSe2/MoS2 stacked channel FET is integrated into an OLED circuit. Figure 4a shows an OM image of the HS memory FET used for this demonstration, and its transfer curves are displayed in Fig. 4b (more details for the device are found along with output characteristics in Supplementary Fig. 5). Figure 4b also presents the OLED pixel circuit in the inset, where a green OLED device is connected in series. Under a supplied voltage (VDD) and 1 s VGS pulse, OLED current (IOLED) was controlled. Figure 4c shows the retention properties of IOLED after VGS pulses for OLED ON and OFF states. An ON/OFF IOLED ratio of ~200 seems maintained for longer than 1500 s. Inset OLED pixel images at different retentions of 500 s, 1000 s, and 1500 s, show almost the same brightness. d Time-domain IOLED plot and corresponding VOUT – time plot, as obtained under different pulses. Gray-scale pixel illumination appears clearly with five brightness levels as shown in the inset green OLED images.

**Synaptic memory behavior of stack channel FET**

Apart from the VGS-modulated multilevel memory functions, we have also attempted to mimic the biological synapse using our MoSe2/MoS2 stacked channel FET as inspired by cyclic memory endurance experiments. In the biological nervous system\(^2\), the neurons are connected to each other through the synapse, which conveys electrical or chemical signals from pre-neuron to post-neuron. And the connectivity between neurons, named synaptic weight, can be modulated by the number of released neurotransmitters as shown in Fig. 5a. In our stack (or synaptic stack) channel FET of Fig. 5a, constant VGS in pulse series serves as a presynaptic voltage (VPre) spike, which leads to a change of synaptic weight. A pulse (spike) number-modulated channel conductance, G (defined as ID divided by VDS), is matched with synaptic weight, where ID is analogous to postsynaptic current (Ips). We applied 100 excitatory pulses for potentiation (negative VPre) and another 100 inhibitory pulses for depression (positive VPre) in series, as depicted in Fig. 5b. Figure 5c shows the monitored P–D plots depending on the diverse amplitude of excitatory pulse (−3.6 V, −3.7 V, −4 V, and −4.3 V with 60 μs width) and fixed inhibitory pulse amplitude of 3.7 V with 60 μs width; we only vary the amplitude of an excitatory pulse to analyze the variation behavior of maximum/minimum channel conductance ratio (GMax/GMin). Here, GMax appears to become large by increasing the amplitude of pulses; a larger magnitude of VPre pulse evacuates more trapped electrons from the MoSe2/MoS2 interfacial trap DOS, resulting in larger channel conductance ratio. Apart from the VGS-modulated multilevel memory functions, we have also attempted to mimic the biological synapse using our MoSe2/MoS2 stacked channel FET as inspired by cyclic memory endurance experiments. In the biological nervous system\(^2\), the neurons are connected to each other through the synapse, which conveys electrical or chemical signals from pre-neuron to post-neuron. And the connectivity between neurons, named synaptic weight, can be modulated by the number of released neurotransmitters as shown in Fig. 5a. In our stack (or synaptic stack) channel FET of Fig. 5a, constant VGS in pulse series serves as a presynaptic voltage (VPre) spike, which leads to a change of synaptic weight. A pulse (spike) number-modulated channel conductance, G (defined as ID divided by VDS), is matched with synaptic weight, where ID is analogous to postsynaptic current (Ips). We applied 100 excitatory pulses for potentiation (negative VPre) and another 100 inhibitory pulses for depression (positive VPre) in series, as depicted in Fig. 5b. Figure 5c shows the monitored P–D plots depending on the diverse amplitude of excitatory pulse (−3.6 V, −3.7 V, −4 V, and −4.3 V with 60 μs width) and fixed inhibitory pulse amplitude of 3.7 V with 60 μs width; we only vary the amplitude of an excitatory pulse to analyze the variation behavior of maximum/minimum channel conductance ratio (GMax/GMin). Here, GMax appears to become large by increasing the amplitude of pulses; a larger magnitude of VPre pulse evacuates more trapped electrons from the MoSe2/MoS2 interfacial trap DOS, resulting in larger channel conductance of synaptic device. Figure 5d shows two P–D pulse combinations (−4.3 V/3.7 V and −3.7 V/3.7 V). In fact, a high GMax/GMin is no more important, once the ratio attains a certain high value (over ~40)\(^3\). The linearity of P–D plots becomes rather important in this case\(^4\). According to the P–D plots of Fig. 5d, the potentiation linearity of the low GMax/GMin ratio case is apparently better than...
that of the high ratio plot. On the one hand, depression plots of both cases appear much deviated from ideal linearity. The very different linearity and asymmetry between potentiation and depression plot are intrinsically attributed to the band offset in MoSe2/MoS2 heterojunction band diagram, as shown in Fig. 2b. Owing to the shape of band offset at the heterojunction, electron trapping (for depression) is much easier than de-trapping (potentiation). That is why the linearity of potentiation is better than that of depression, which also causes asymmetric P-D curves. More details about the dependency of linearity on pulse amplitude are included in Supplementary Fig. 7 and Supplementary Table 2. We eventually perform ANN simulation for three cases: assuming perfect linearity and using actual data in any P-D plots of Fig. 5c, d. Figure 5e shows the schematics of multilayer perceptron (MLP) neural network constructed for classification of MNIST handwritten digits. It has three layers of neurons, and those layers are composed of 784 input neurons, 100 hidden neurons, and 10 output neurons, respectively. Three neuronal layers are connected by modeled synapse on the basis of the P-D plot of our stack channel memory. The modeled synapse is composed of our stack channel FET and switching transistor as shown in a circuit diagram of our artificial neural network (ANN) composed of modeled synapse and peripheral circuits, to perform program and read processes. The synapse consists of our stack channel FET and a switching transistor. g Simulation results exhibiting that the ANN using our stack channel FET demonstrates ~94% accuracy on average by assuming perfect linearity but only ~77% accuracy from actual P-D behavior whether GMax/GMin ratio is 115 or 61.7.

In summary, we have studied multilevel memory devices operating under short pulse low voltage conditions (60 μs shortest, 7 V maximum), based on a vdW heterostacked n-MoSe2/n-MoS2 channel FET. Our HS channel FET initially appeared to operate as a simple bistable nonvolatile memory FET with Program and Erase states. The HS memory device properly exploited the VSS-induced trapping/de-trapping phenomena for such Program/Erase functioning, which was nicely maintained for 10^5 s-long retention times due to the existence of a heterojunction energy barrier between MoS2 and MoSe2. More interestingly, trapped electron density could be incrementally modulated by the magnitude of pulsed VGS, enabling the HS device to achieve multilevel long-term memory. For a practical multiscale display demonstration, five different levels of IC were visualized with modulated light emissions when our memory FET was integrated into a green OLED pixel circuit. In addition, our device was applied to a synaptic memory function. Based on the P-D characteristics of our device under multiple 60-μs-short VGS pulses, the simulation resulted in an average ~94% recognition accuracy. We conclude that our HS channel FETs are interesting and promising enough to cope with future demanding multilevel memory electronics for the fourth industrial revolution.
METHODS
Device fabrication
Glass substrates (Eagle 2000) were cleaned by sonication for 30 min in acetone and ethyl alcohol, respectively. As a bottom gate electrode, Au/Ti (10/5 nm) was patterned by photolithography, DC sputter-deposition, and the lift-off process. A 30-nm-thick Al2O3 was deposited by atomic layer deposition at 110 °C as a gate dielectric insulator. For the ultrathin PS-brush layer on Al2O3, dimethyl-chlorosilane-terminated PS (Polymer Source, Product No. P3881-SSiCl) was dissolved in toluene (Aldrich) solvent at a 10 mg/mL ratio. Before the PS-brush coating, oxygen plasma was conducted on the Al2O3 surface to make enough hydroxyl groups. PS solution was spin-coated and then the substrate was put in the oven to be annealed at 170 °C for 48 h. During annealing, an ultrathin PS-brush layer was formed making a covalent bond with hydroxyl groups on the surface of Al2O3. The unreacted portion of PS was removed by soaking in toluene solvent. Then the PS-brush treated substrates were again annealed for 24 h in a vacuum oven. Mechanically exfoliated MoS2, MoSe2 flakes were transferred in sequence on a specific position of the PS-brush/Al2O3, where the patterned bottom gate was located. Finally, we deposited and patterned Au (70 nm) as the source and drain electrode by the same processes used for the gate electrode. Here, Au was chosen in consideration of its long-term stability although Ti contact would be better in charge injection as shown in Supplementary Fig. 8.

Device measurements
Basic electrical measurements of our devices were performed with a semiconductor parameter analyzer (Agilent 4155 C) in the dark at room temperature. The capacitances of dielectric materials were measured by a precision LCR meter (Agilent 4284 A) under the same conditions. For the multilevel long-term memory and the artifical synaptic device measurements, a pulse generator (81104 A, Keysight) was used to apply electrical pulses with various amplitudes, widths, and time intervals.

Materials characterization
Cross-sectional STEM samples were prepared by using a focused ion beam (FEI Helios 650 dual beam) for STEM and EDS measurements. Their STEM images and EDS mapping were obtained by transmission electron microscopes (JEOL ARM-200F equipped with image and probe aberration correctors) at 200 kV. Raman spectra were obtained with a 532-nm wavelength laser (LabRam Aramis, Horiba Jovin Yvon). A scanning probe microscopy (NX-10, Park system) was used for SKPM measurements.

Array simulation
“NeuroSim+” simulator was utilized for the ANN simulation. An MLP neural network was composed of 784 input neurons, 100 hidden neurons, and 10 output neurons. The input layer of 784 neurons corresponded to the postprocessing 28×28 pixels of an MNIST image. Ten neurons of the output layer matched with ten classes of digits from 0 to 9. For the simulation of recognition accuracy calculations, the SGD algorithm was applied for training based on the potentiation-depression plots of our stack channel FET characteristics.

DATA AVAILABILITY
The authors confirm that the data supporting the findings of this study are available within the article and its supplementary information.

Received: 7 October 2021; Accepted: 9 February 2022; Published online: 21 March 2022

REFERENCES
1. Turing, A. M. Computing machinery and intelligence, Parsing the Turing test (Springer, Dordrecht, 2009).
2. Copeland, B. J. The Essential Turing (Clarendon Press, Oxford, 2004).
3. Prezioso, M. et al. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature 521, 61–64 (2015).
4. Choi, S., Shin, J. H., Lee, J., Sheridan, P. & Lu, W. D. Experimental demonstration of feature extraction and dimensionality reduction using memristor networks. Nano Lett. 17, 3113–3118 (2017).
5. Zhao, M., Gao, B., Tang, J., Qian, H. & Wu, H. Reliability of analog resistive switching memory for neuromorphic computing. Appl. Phys. Rev. 7, 011301 (2020).
6. Gkoupidenis, P., Koutsouras, D. A. & Malliaras, G. G. Neuromorphic device architectures with global connectivity through electrolyte gating. Nat. Commun. 8, 1–8 (2017).
7. Song, K. M. et al. Skyrmion-based artificial synapses for neuromorphic computing. Nat. Electron. 3, 148–155 (2020).
8. Ge, C. et al. A ferrite synaptic transistor with topotactic transformation. Adv. Mater. 31, 1900379 (2019).
9. Novoselov, K., Mishchenko, O. A., Carvalho, O. A. & Neto, A. C. 2D materials and van der Waals heterostructures. Science 335, aaq4939 (2016).
10. Wang, H. et al. Two-dimensional heterostructures: fabrication, characterization, and application. Nanoletters 6, 12250–12272 (2014).
11. Rhodes, D., Chae, S. H., Ribeiro-Palau, R. & Hone, J. Disorder in van der Waals heterostructures of 2D materials. Nat. Mater. 18, 541–549 (2019).
12. Sebastian, A., Pannone, A., Radhakrishnan, S. S. & Das, S. Gaussian synapses for probabilistic neural networks. Nat. Commun. 10, 1–11 (2019).
13. Bertolazzi, S., Krasnokhodov, D. & K. A. Nonvolatile memory cells based on MoS2/graphene heterostructures. ACS Nano 7, 3246–3252 (2013).
14. Lee, Y. T. et al. Nonvolatile charge injection memory based on black phosphorous 2D nanosheets for charge trapping and active channel layers. Adv. Funct. Mater. 26, 5701–5707 (2016).
15. Li, D., Chen, M., Zong, Q. & Zhang, Z. Floating-gate manipulated graphene-black phosphorus heterojunction for nonvolatile ambipolar schottky junction memories, memory inverter circuits, and logic rectifiers. Nano Lett. 17, 6353–6359 (2017).
16. Choi, M. S. et al. Controlled charge trapping by molybdenum disulphide and graphene in ultrathin heterostructured memory devices. Nat. Commun. 4, 1–7 (2013).
17. Park, S. et al. Nonvolatile and neuromorphic memory devices using interfacial traps in two-dimensional WSe2/MoTe2 stack channel. ACS Nano 14, 12064–12071 (2020).
18. Lee, D. et al. Multibit MoS2 photocurrent memory with ultrahigh sensitivity. Adv. Mater. 28, 9196–9202 (2016).
19. Tran, M. D. et al. Two-terminal multibit optical memory via van der Waals heterostructure. Adv. Mater. 31, 1807075 (2019).
20. Chen, M. et al. Multibit data storage states formed in plasma-treated MoS2 transistors. ACS Nano 8, 4023–4032 (2014).
21. Lee, J. et al. Monolayer optical memory cells based on artificial trap-mediated charge storage and release. Nat. Commun. 8, 1–8 (2017).
22. Wang, Y. et al. Band-tailored van der Waals heterostructure for multilevel memory and artificial synapse. InfoMat 3, 917–928 (2021).
23. Jeong, Y. et al. 2D MoSe2 transistor with polymer-brush/channel interface. Adv. Mater. Interfaces 5, 1800812 (2018).
24. Jeong, Y. et al. Integrated advanced technologies from perovskite photovoltaic cell and 2D MoTe2 transistor towards self-power energy harvesting and photosensing. Nano Energy 63, 103833 (2019).
25. Kim, M., Anjum, M. A. R., Lee, M., Lee, B. J. & Lee, J. S. Activating MoS2 basal plane with Ni2P nanoparticles for hydrogen evolution reaction in acidic media. Adv. Funct. Mater. 29, 1809151 (2019).
26. Rho, Y. et al. High-speed direct writing of MoS2 by maskless and gas-free laser-assisted selenization process. J. Phys. Chem. C 124, 19333–19339 (2020).
27. Berg, M. et al. Layer dependence of the electronic band alignment of few-layer MoTc2 single layers. ACS Energy Lett. 6353 (2020).
28. Choi, S. et al. SiGe epitaxial memory for neuromorphic computing with reproducible high performance based on engineered dislocations. Nat. Mater. 17, 335–340 (2018).
29. Ham, S. et al. One-dimensional organic artificial multi-synapses enabling electronic textile neural network for wearable neuromorphic applications. Sci. Adv. 6, eaba1178 (2020).
30. Jang, J.-W. et al. Optimization of conductance change in Pr3+-Ca,MnO2-based synaptic devices for neuromorphic systems. IEEE Electron Device Lett. 36, 457 (2015).
31. Ahmed, T. et al. Fully Light-controlled memory and neuromorphic computation in layered black phosphorus. Adv. Mater. 33, e2004207 (2020).

Published in partnership with FCT NOVA with the support of E-MRS

npj 2D Materials and Applications (2022) 23
ACKNOWLEDGEMENTS
The authors acknowledge the financial support from the National Research Foundation of Korea (SRC program: grant no. 2017R1A5A1014862, vdWMRC. The authors in KAIST were funded from the NRF (grant no. 2019M3F3A1A01074452, 2019M3F3A1A01074452, and 2020M3F3A2A01085755). This work was supported by Yonsei Signature Research Cluster Program in the year 2021.

AUTHOR CONTRIBUTIONS
Y.J. and H.J.L. contributed equally to this work. Y.J. and S.I. conceived and designed the study. Y.J. and H.J.L. fabricated samples and carried out experiments under the guidance of S.I. and with the help of the other authors. J.P. and S.C. conducted ANN simulation and analyzed the synaptic behavior of the device. S.L. and K.K. executed the STEM experiment and analyzed EDS data. T.K. helped with the electrical measurement of the memory device. All authors contributed to the discussion of this work. Y.J. and S.I. wrote the manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41699-022-00295-8.

Correspondence and requests for materials should be addressed to Seongil Im.

Reprints and permission information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022