

# 3D-integrated multilayered physical reservoir array for learning and forecasting time-series information

Received: 29 November 2023

Accepted: 22 February 2024

Published online: 06 March 2024

 Check for updatesSanghyeon Choi<sup>1,2,6</sup>, Jaeho Shin<sup>1,3</sup>, Gwanyeong Park<sup>1</sup>, Jung Sun Eo<sup>1</sup>, Jington Jang<sup>1,7</sup>, J. Joshua Yang<sup>2</sup>  & Gunuk Wang<sup>1,4,5</sup> 


A wide reservoir computing system is an advanced architecture composed of multiple reservoir layers in parallel, which enables more complex and diverse internal dynamics for multiple time-series information processing. However, its hardware implementation has not yet been realized due to the lack of a high-performance physical reservoir and the complexity of fabricating multiple stacks. Here, we achieve a proof-of-principle demonstration of such hardware made of a multilayered three-dimensional stacked  $3 \times 10 \times 10$  tungsten oxide memristive crossbar array, with which we further realize a wide physical reservoir computing for efficient learning and forecasting of multiple time-series data. Because a three-layer structure allows the seamless and effective extraction of intricate three-dimensional local features produced by various temporal inputs, it can readily outperform two-dimensional based approaches extensively studied previously. Our demonstration paves the way for wide physical reservoir computing systems capable of efficiently processing multiple dynamic time-series information.

Big data has emerged as a critical driving force to the data-driven industrial revolution, which underpins decision-making processes in modern society, enhances efficiency, and promotes innovation<sup>1,2</sup>. The ability to collect, analyze, and interpret big data supports companies and societies involved in artificial intelligence in maintaining a competing edge. Among the various types of data, time-series data, which are sequential observations collected over specific time frames, have become increasingly important as dynamic information in applications such as biometric analysis<sup>3</sup>, weather forecasting<sup>4</sup>, stock chart estimation<sup>5</sup>, and water inflow prediction<sup>6</sup>. Their importance arises from the inherently ordered nature, which allows predicting future

behaviors in dynamic information using estimated models that describe correlated features.

In artificial intelligence, time-series data have been extensively explored through approaches including network, statistical, and transformer models<sup>7,8</sup>. A promising approach is the implementation of reservoir computing (RC) based on nonlinear dynamic systems such as physical reservoirs (PRs) to replace computationally intensive software-based methods<sup>9,10</sup>. Remarkably, memristive systems feature inherent nonlinearity and short-term memory capability, which enable a single memristor node to replace a complex nonlinearly coupled loop formed by multiple nodes in a software reservoir<sup>11–18</sup>. Since the

<sup>1</sup>KU-KIST Graduate School of Converging Science and Technology, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea.

<sup>2</sup>Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089, USA. <sup>3</sup>Department of Chemistry, Rice University, 6100 Main Street, Houston, TX 77005, USA. <sup>4</sup>Department of Integrative Energy Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea. <sup>5</sup>Center for Neuromorphic Engineering, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea. <sup>6</sup>Present address: Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA. <sup>7</sup>Present address: School of Computer and Information Engineering, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Republic of Korea.  e-mail: [jjoshuay@usc.edu](mailto:jjoshuay@usc.edu); [gunukwang@korea.ac.kr](mailto:gunukwang@korea.ac.kr)

first demonstration of memristive physical RC<sup>11</sup>, several approaches based on memristors, such as the virtual node concept<sup>14</sup>, self-organized nanowire network<sup>17</sup>, input masking<sup>16</sup>, and a fully analog framework<sup>18</sup>, have been suggested to date. Although these proposed physical RC systems are noteworthy, most of them are fabricated on single-stacked and two-dimensional (2D) memristive device frameworks and inevitably restricted to a single reservoir architecture. The 2D-based approaches eventually encounter physical limitations in addressing multi-variant time information. In addition, several multi-layered physical RC systems have been recently proposed<sup>19,20</sup>. These systems demonstrated the promising aspects of the advanced reservoir architecture, including improved reservoir capacity and optimization. However, these approaches have a 2D lateral architecture or rely on the input design rule, which requires additional multilayered PR architecture capable of parallel processing, simplification, and versatility to efficiently address multi-variable temporal inputs (detailed in Supplementary Note 1). Hence, it is desirable to realize a wide physical RC with multiple-stacked PRs based on three-dimensional (3D) integrated memristive architecture. This 3D approach is anticipated to enable the extraction of richer features and multiple local traits from temporal inputs simultaneously, performing a more accurate, reliable, and rapid RC than the 2D-based single physical RC. In addition, the multiple-stacked PRs can provide signal mapping capabilities with a smaller footprint than a single PR. Consequently, 3D-integrated memristive PRs architecture is a significant step in the development of a next-generation physical RC. However, the achievement of the wide RC based on multiple-stacked PRs remains elusive within existing memristive frameworks as well as other device platforms due to the complex fabrication steps for 3D stacked structures and the stringent requirements on the PR properties.

For instance, many memristors show unreliable switching characteristics such as variation, instability, and low yield<sup>21,22</sup> owing to the intrinsic nature of switching filaments, which makes it extremely challenging to meet the high reproducibility requirement of PR states from cycle to cycle and from device to device. Additionally, the matrix array structure demands additional components, such as transistors or selectors, to ensure reliable programming and reading by preventing crosstalk<sup>23,24</sup>. This leads to complex device structure and fabrication process (i.e., incorporating 1 T) as well as increased power consumption and active node size. Moreover, memristors at matrix nodes must operate in a low programming voltage range because time-series data processing requires a considerable number of input pulses with varying intervals ( $\Delta t$ ) over a specific time frame for low-power operation<sup>14,15</sup>. Furthermore, addressing and predicting correlations between various dynamic information sources require the ability to handle multiple time series concurrently in each time frame without mutual interference<sup>25–27</sup>. In other words, many independent reservoir states controlled by sequential input pulses with varying  $\Delta t$  are required for multiple dynamic tasks to be performed simultaneously without interference. To effectively tackle these challenges all at once, we propose and demonstrate the wide RC based on a multilayered 3D-integrated memristive array structure with several independent reservoir states at each matrix layer, crosstalk-free and robust switching characteristics, high device yield, and low programming voltages. In other words, the primary objective of our research is to design, fabricate, and implement a wide physical RC platform, enabling its capacity to be widely applied in sophisticated reservoir configurations and providing functional benefits that go beyond simply increasing the density of devices.

Our approach can efficiently capture and process 3D local dynamic features of multiple time series by implementing a 3D stacked  $3 \times 10 \times 10$  tungsten oxide ( $\text{WO}_x$ ) array architecture. Each memristor cell is designed based on asymmetric interfacial Schottky barrier engineering to achieve crosstalk-free of signals throughout the array. The 3D stacked array exhibits highly reliable and robust operation with

characteristics such as a high yield (>98%), exceptional stability (>4000 consecutive sweeps and > $10^5$  cycles), switching uniformity (-0.24%), and low operating voltage (-0.7 V). We demonstrate reliable reservoir operations under various time-dependent electrical inputs without overlapping reservoir states. As a proof of concept, we perform spatiotemporal pattern classification of biological cell positions and behavior prediction of a time-dependent chaotic Lorenz attractor. Our proposal achieves a higher and faster (~4.5 times) classification accuracy for 3D positions of a biological cell using 60% fewer memristor cells, and a tenfold reduction in prediction error for the chaotic Lorenz attractor compared with the use of a single reservoir. The proposed 3D physical RC architecture supported by reliable memristors opens an avenue for efficient and compact physical structure for reservoir computing capable of time-series processing in artificial intelligence systems.

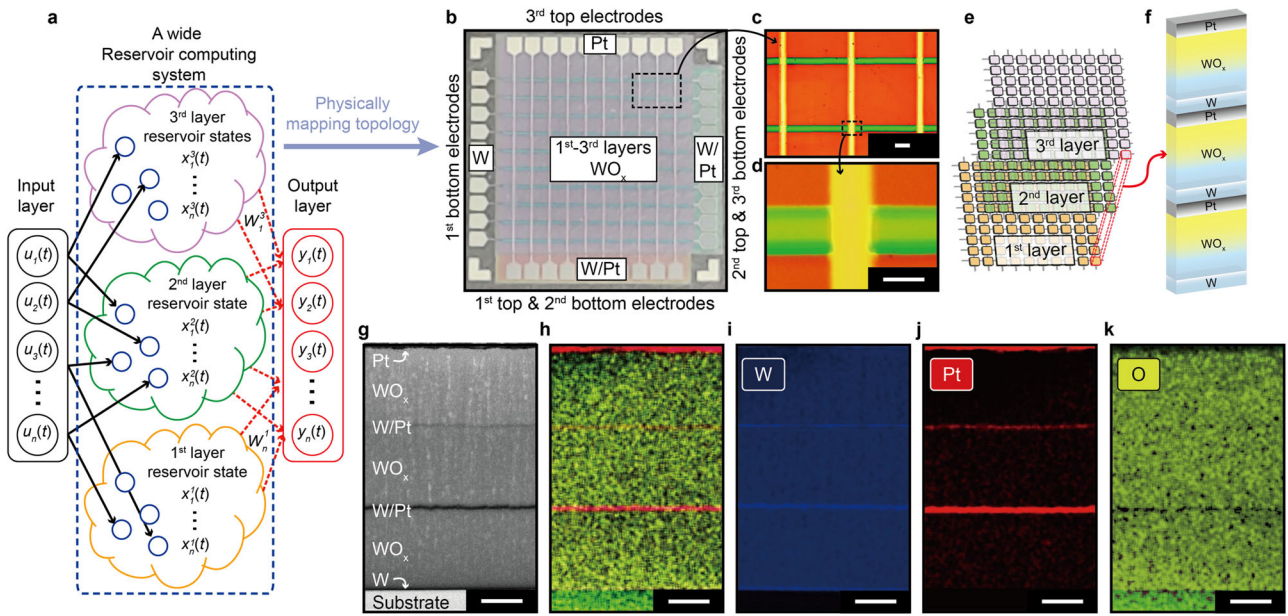
## Results

### Three-layer stacked physical reservoirs

Figure 1a illustrates a wide RC system consisting of three reservoir layers (yellow, green, and purple) and an output layer. These reservoirs can perform a nonlinear projection of temporal inputs ( $u_n(t)$ ) onto reservoir states ( $x_n^i(t)$ ) per feature space via the different nodes of the reservoirs (blue circles), where  $i$  and  $n$  are the order of the reservoir layer ( $i = 1, 2, 3$ ) and number of nodes ( $n = 1, 2, \dots, n$ ), respectively<sup>9,10</sup>. As the collective reservoir states are delivered from each reservoir layer to the output layer, the weights ( $W_n^i$ ) between them can be iteratively updated for the desired outputs ( $y_n(t)$ ) according to an algorithm (e.g., backpropagation, linear regression). Such learning process through the output layer can be readily executed based on a linear combination of the weighted reservoir states, thereby reducing the learning costs when compared with other time-series processing algorithms. Moreover, it has been demonstrated through an algorithmic and theoretical approach that the utilization of a wide RC system is more effective in handling multiple dynamic time information simultaneously by extracting rich local features through multiple independent reservoir layers, in comparison to a single RC system (Supplementary Fig. 1 and Supplementary Note 2)<sup>25–27</sup>. As shown in Figs. 1a, b, the wide RC system can be physically mapped onto a 3D stacked three-layer memristive crossbar array, where each layer acts as a reservoir (three reservoirs in total). Figure 1b shows a top-view optical microscopy image of the fabricated 3D stacked three-layer  $3 \times 10 \times 10$  memristive crossbar array, where each layer consists of Pt/ $\text{WO}_x$ /W memristor cells at each crosspoint (300 cells). In Fig. 1b, the first, second, and third  $\text{WO}_x$  layers are indicated in orange, green, and purple, respectively, confirming the formation of a 3D stacked array. Figures 1c, d show the 3D stacked  $3 \times 2 \times 3$  subarray (Fig. 1c) with a magnified view of the vertically integrated cells at a crosspoint (Fig. 1d), revealing a cell line width of approximately 100  $\mu\text{m}$ . To facilitate the understanding of the 3D stacked architecture, the corresponding schematics are shown in Figs. 1e, f. Array fabrication with a cell line width of 20  $\mu\text{m}$  is also demonstrated using conventional photolithography (Supplementary Fig. 2). Moreover, similar switching feature at nanoscale is demonstrated, as detailed below. Figures 1g–k show cross-sectional high-resolution transmission electron microscopy (HR-TEM) images and the corresponding energy dispersive spectroscopy (EDS) results of the 3D stacked array at a crosspoint, verifying the well-defined vertical integration of the three Pt/ $\text{WO}_x$ /W junction structures. The detailed fabrication processes are provided in the Supplementary Information (Supplementary Fig. 3) and Methods section. Note that the benefits of our approach are further discussed in Supplementary Figs. 4, 5, Supplementary Table 1, and Supplementary Note 3, 4 of Supplementary Information.

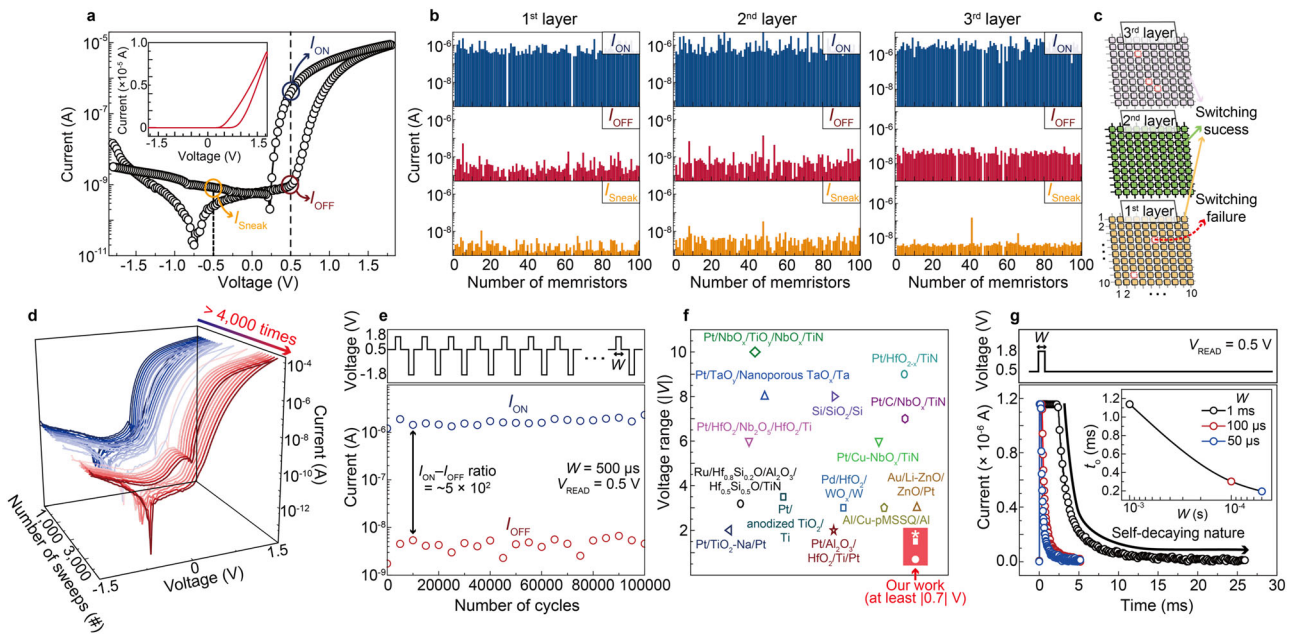
### Electrical characteristics of the 3D physical reservoirs

Figure 2a shows a representative self-rectifying current–voltage ( $I$ – $V$ ) switching curve of a memristor in the 3D stacked  $\text{WO}_x$  PRs. The  $I_{\text{ON}}/I_{\text{OFF}}$



**Fig. 1 | 3D stacked  $WO_x$  physical reservoir (PR) arrays.** **a** Schematic of a wide reservoir computing (RC) system with multiple reservoirs (see Supplementary Note 2). **b** Top-view optical image of fabricated 3D stacked three-layer  $3 \times 10 \times 10$  crossbar array with vertical integration of three Pt/ $WO_x$ /W memristors at each crosspoint. The first (1st), second (2nd), and third (3rd)  $WO_x$  layers are shown in orange, green, and purple, respectively. The fabrication of each layer is described in the Supplementary Information (Supplementary Fig. 3). The 3D stacked array can naturally map a wide RC system onto hardware. **c**, **d** Top-view optical image of  $3 \times 3 \times 3$  sub-array (**c**) and magnified view at crosspoint (**d**) with line width of

100  $\mu\text{m}$ . The scale bar for the top and bottom image is 200  $\mu\text{m}$  and 100  $\mu\text{m}$ , respectively. **e** Schematic of fabricated 3D stacked three-layer memristive crossbar array. Each memristor cell located at the first, second, and third layers is shown in orange, green, and purple, respectively. **f** Magnified schematic of vertical integration of three Pt/ $WO_x$ /W memristors at a certain crosspoint ( $3 \times 1 \times 1$ ) in panel (**e**) (red dotted line). **g–k** Cross-sectional high-resolution transmission electron microscopy (HR-TEM) image of three stacked memristors at crosspoint (**g**) and its corresponding energy dispersive spectroscopy (EDS) results for all (**h**) and each (**i–k**) of its elements. The scale bar for all (**g–k**) is 100 nm.



**Fig. 2 | Electrical characteristics of 3D Pt/ $WO_x$ /W memristor array.**

**a** Representative  $I$ - $V$  curve of Pt/ $WO_x$ /W memristor without initial forming and exhibiting self-rectifying bipolar switching. The top Pt electrode and bottom W electrode were biased and grounded, respectively. The inset shows the same  $I$ - $V$  curve on a linear scale. **b** Statistical histograms of switching parameters  $I_{ON}$ ,  $I_{OFF}$ , and  $I_{Sneak}$  for all the memristors from the first to third layers (300 memristors). **c** Schematic of 3D stacked  $WO_x$  array illustrating switching success and failure (i.e., short circuit) as colored and red dotted boxes, respectively. **d** Operation under

consecutive  $I$ - $V$  sweeps over 4000 times for  $WO_x$  memristor in 3D stacked array. **e** Endurance cycling test of  $WO_x$  memristor over  $10^5$  cycles by programming at 1.8 V and  $-1.8$  V for  $W = 500 \mu\text{s}$ .  $I_{ON}$  and  $I_{OFF}$  were read at  $V_{READ} = 0.5$  V. **f** Operating voltage range  $|V|$  for  $WO_x$  memristor (red box) and other self-rectifying memristors. **g** Dynamic  $I$  behavior over time  $t$  for  $WO_x$  memristor after applying voltage pulses of 1.8 V for  $W = 1$  ms, 100  $\mu\text{s}$ , and 50  $\mu\text{s}$ , demonstrating a self-decaying nature (i.e., short-term memory) with different  $t_0$  (inset).



and rectification ( $I_{\text{ON}}/I_{\text{Sneak}}$ ) ratios are approximately  $4.1 \times 10^2$  and  $4.8 \times 10^2$ , respectively, where  $I_{\text{ON}}$  and  $I_{\text{OFF}}$  are the ON and OFF currents at reading voltage  $V_{\text{READ}} = 0.5 \text{ V}$ , respectively, and  $I_{\text{Sneak}}$  is the sneak current at  $-V_{\text{READ}}$ . The programming voltage ranges in  $\pm 1.8 \text{ V}$ . An initial electroforming step is not required for the fabricated array memristors. This switching phenomenon is attributed to the robust Schottky barrier formed at the Pt/ $\text{WO}_x$  interface and dynamic modulation of the Schottky barrier at the  $\text{WO}_x/\text{W}$  interface, which is controlled by the electric-field-driven migration of oxygen vacancies<sup>28,29</sup>. It can be also observed even in the absence of an applied compliance current (Fig. 2a). This operational advantage can reduce the burden on peripheral circuitry for strict current control, hence, it enables our 3D stacking array to be denser and more scalable. Note that the shift of voltage position for minimum currents in Fig. 2a might be associated with the transient formation of an internal electric field<sup>30–32</sup> (Supplementary Note 5). Figure 2b shows the statistical histograms of  $I_{\text{ON}}$ ,  $I_{\text{OFF}}$ , and  $I_{\text{Sneak}}$  for all the memristors from the first to third layers of the 3D stacked  $\text{WO}_x$  PRs (300 memristors). The  $I$ - $V$  switching curves of all the reservoir layers are provided in the Supplementary Information (Supplementary Figs. 6–8). Similar  $I$ - $V$  switching curves are observed regardless of the location of the reservoir layers. The average values for  $I_{\text{ON}}$ ,  $I_{\text{OFF}}$ , and  $I_{\text{Sneak}}$  are  $(1.40 \pm 0.10) \times 10^{-6} \text{ A}$ ,  $(4.05 \pm 0.38) \times 10^{-9} \text{ A}$ , and  $(2.23 \pm 0.16) \times 10^{-9} \text{ A}$ , respectively. All the distributions of  $I_{\text{ON}}$ ,  $I_{\text{OFF}}$ , and  $I_{\text{Sneak}}$  are well-fitted by the lognormal distribution curve (Supplementary Fig. 9). Hence, well-defined switching parameters with acceptably low variations are observed in the 3D stacked  $\text{WO}_x$  PRs. The device yield of the layers is 98.3% (97, 100, and 98/100 for the first, second, and third layers, respectively), as illustrated in the schematic of Fig. 2c (see Supplementary Figs. 6–8). These results support the reproducibility and expandability of the 3D stacked  $\text{WO}_x$  PRs for large-scale physical RC systems.

Figure 2d, e show the consecutive switching curves for more than 4000 times and over  $10^5$  cycles with an  $I_{\text{ON}}/I_{\text{OFF}}$  ratio of approximately  $5 \times 10^2$ , respectively, demonstrating an excellent operating stability with average cycle variation of 0.24% (Supplementary Fig. 10). In addition, the multiple 3D stacked  $\text{WO}_x$  PRs exhibit the lowest operating voltage range (at least  $|0.7| \text{ V}$ , Fig. 2f and Supplementary Fig. 11) with robust self-rectification (Supplementary Fig. 12), indicating the possibility of lower power consumption than existing self-rectifying memristors<sup>31,33–45</sup>. Figure 2g shows the current responses of the memristor cell over time after applying single voltage pulses of different widths ( $W$ ). The current abruptly increases close to  $I_{\text{ON}}$  soon after applying a single voltage pulse and then decays spontaneously and exponentially over time ( $t$ ) under  $V_{\text{READ}}$  (i.e.,  $I \propto \exp(-t/t_0)$ , where  $t_0$  is the characteristic decay time), eventually returning to  $I_{\text{OFF}}$ . This behavior was still observed to persist despite the application of strong programming voltage scheme in both air and vacuum environments (Supplementary Fig. 13). This spontaneous decay in the switching current cannot be attributed to the protonation-assisted switching mechanism, as the metallic  $\text{H}_x\text{WO}_3$  phase is absent and the switching occurs over a short period of time<sup>46</sup>. This self-decaying nature (i.e., short-term memory) can describe time-dependent internal states that enable temporal processing. Because width  $W$  can control  $t_0$ , a 3D stacked  $\text{WO}_x$  memristor cell can process time-series data over varying time frames. Note that  $t_0$  was observed to decrease as the  $W$  decreased based on a power function (Supplementary Fig. 14). Consequently, as compared with other reported memristor devices<sup>47–50</sup>, our proposed  $\text{WO}_x$  memristor can exhibit highly reliable selector-less short-term switching performances, operational advantages, and scalability for showcasing the 3D-integrated multilayered PR arrays.

### Switching mechanism of the 3D physical reservoirs

The interfacial barrier heights were investigated according to the  $I_{\text{ON}}$  and  $I_{\text{OFF}}$  states to understand the self-rectifying switching behavior of the 3D stacked  $\text{WO}_x$  memristor cell. Initially, the Pt/ $\text{WO}_x/\text{W}$  memristor

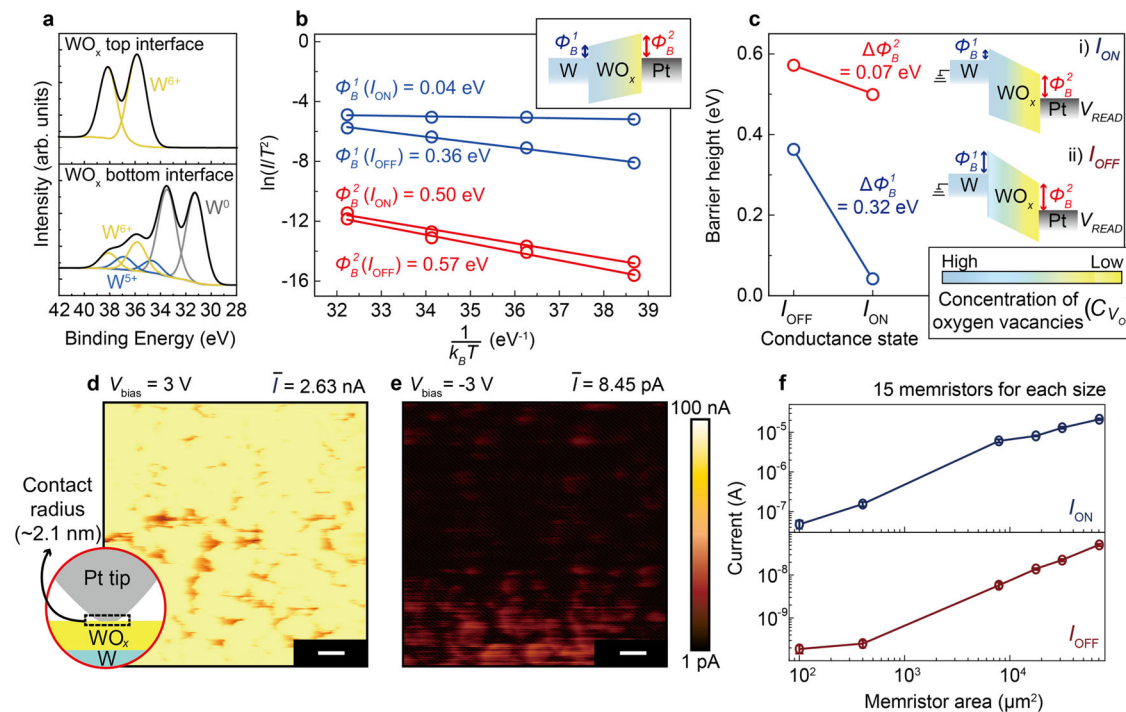
was designed based on the asymmetric distribution of oxygen vacancy ( $V_{\text{O}}$ ) in  $\text{WO}_x$  between Pt and W electrodes. Figure 3a shows the ex-situ X-ray photoelectron spectroscopy (XPS) results with timed  $\text{Ar}^+$  bombardment for the W ion peaks at the top and bottom  $\text{WO}_x$  interfaces. Compared with the top  $\text{WO}_x$  interface, there are two peaks corresponding to  $\text{W}^{5+}$  and  $\text{W}^0$  at the bottom  $\text{WO}_x$  interface (top and bottom of Fig. 3a)<sup>51,52</sup>. This indicates the enrichment of  $V_{\text{O}}$  in the direction of the bottom  $\text{WO}_x$  interface, which increases the concentration of  $V_{\text{O}}$  ( $C_{V_{\text{O}}}$ ) at the  $\text{WO}_x/\text{W}$  interface. Because the Fermi level of  $\text{WO}_x$  increases (decreases) with high (low)  $C_{V_{\text{O}}}$ <sup>53</sup>, asymmetric barrier heights  $\Phi_B^1$  and  $\Phi_B^2$  at the  $\text{WO}_x/\text{W}$  and Pt/ $\text{WO}_x$  interfaces can be formed. A detailed discussion and additional XPS results are presented in Supplementary Note 5 and Supplementary Figs. 15, 16. The combined effects of the sputtering parameters and electrodes were also investigated (see the Methods section and Supplementary Figs. 17–19). The structural analysis results are consistent with the visual inspection of  $\text{WO}_x$  (Supplementary Fig. 20).

Figure 3b shows the temperature ( $T$ )-dependent  $I_{\text{ON}}$  and  $I_{\text{OFF}}$  states of the Pt/ $\text{WO}_x/\text{W}$  memristor and their interfacial barrier heights ( $\Phi_B^1$  at  $\text{WO}_x/\text{W}$  and  $\Phi_B^2$  at Pt/ $\text{WO}_x$ ) through the Arrhenius plot. Based on the fitting results shown in Fig. 3b (solid lines; see the Methods section),  $\Phi_B^1$  and  $\Phi_B^2$  are shown according to the  $I_{\text{ON}}$  and  $I_{\text{OFF}}$  states in Fig. 3c. The asymmetry between  $\Phi_B^1$  and  $\Phi_B^2$  contributes to the self-rectifying feature shown in Fig. 2a. A larger change in  $\Phi_B^1$  than in  $\Phi_B^2$  is observed during the switching transition, implying that charge transfer through  $\Phi_B^1$  is dominant in determining the switching states. When a positive voltage is applied to the Pt electrode,  $V_{\text{O}}$  migrates downward, and  $C_{V_{\text{O}}}$  increases at the bottom  $\text{WO}_x$  interface<sup>28,29</sup>. In that case,  $\Phi_B^1$  is largely decreased at  $V_{\text{READ}} = 0.5 \text{ V}$ ; then, the conductance state is switched to  $I_{\text{ON}}$ . When a negative voltage is applied to the Pt electrode, the behaviors occur inversely, and the conductance state switches to  $I_{\text{OFF}}$ . Additional experimental results for different thicknesses of  $\text{WO}_x$  under ambient or vacuum environments and the structural analysis of  $\text{WO}_x$  are described in the Supplementary Information (Supplementary Figs. 17, 21–23, Supplementary Note 5). Note that there is a possibility that electron trapping/detrapping process may partially contribute to the switching mechanism, but it is likely to have a minor effect on the switching compared to the electric field-driven  $V_{\text{O}}$  migration (Supplementary Fig. 24 and Supplementary Note 5).

Figures 3d and 3e show maps of the spatial current distribution ( $\bar{I}$ ) on the Pt-tip/ $\text{WO}_x/\text{W}$  junction structure according to the tip bias ( $V_{\text{bias}} = \pm 3 \text{ V}$ ) using conductive atomic force microscopy (c-AFM). The estimated contact radius of the Pt tip is approximately 2.1 nm (Supplementary Fig. 25). Details of the c-AFM measurements are shown in Supplementary Fig. 25 and Methods section. The average values of  $\bar{I}$  are 2.63 nA and 8.45 pA at  $V_{\text{bias}}$  of 3 V and  $-3 \text{ V}$ , respectively, demonstrating a uniform change at the whole scanning area without evident current hotspots (i.e., conductive filaments)<sup>54</sup>. The  $I$ - $V$  switching feature at the nanoscale using c-AFM is consistent with the switching behaviors observed in Fig. 2 (Supplementary Fig. 25). Figure 3f shows statistical  $I_{\text{ON}}$  and  $I_{\text{OFF}}$  according to the memristor size (with a constant  $I_{\text{ON}}/I_{\text{OFF}}$  ratio, Supplementary Fig. 26), demonstrating the homogeneous enrichment and depletion of  $V_{\text{O}}$  near the  $\text{WO}_x$  interfaces<sup>55</sup>. Note that dependency of the  $I_{\text{ON}}/I_{\text{OFF}}$  on the size is almost negligible (Supplementary Fig. 26).

### Physical reservoir computing for multiple time-series data

Temporal processing of multiple dynamic inputs was quantitatively evaluated considering the nonlinear dynamics and short-term memory of 3D stacked  $\text{WO}_x$  PRs (Fig. 4 and Supplementary Figs. 27–30). Figure 4a shows the change in PR states (i.e., current responses) of the memristor cell in response at each timestep ( $t_1$ – $t_4$ ) over a voltage pulse sequence. Pulse sequence [1111] indicates that a voltage pulse ( $V_{\text{INPUT}} = 1.8 \text{ V}$  for 200 ns) is sequentially applied at each timestep ( $t_1$ – $t_4$ ), whereas sequence [0000] indicates that no



**Fig. 3 | Switching mechanism of 3D Pt/WO<sub>x</sub>/W memristor.** **a** W 4f XPS results taken from top and bottom interfaces of WO<sub>x</sub> layer, exhibiting asymmetric profile in concentration of oxygen vacancies ( $C_{V_o}$ ) depending on the position of the WO<sub>x</sub> layer. **b** Arrhenius plot of  $\ln(I/I^2)$  according to  $1/k_B T$  for  $I_{ON}$  and  $I_{OFF}$  states of Pt/WO<sub>x</sub>/W memristor. Variables  $\Phi_B^1$  (blue) and  $\Phi_B^2$  (red) indicate each Schottky barrier height at the W/WO<sub>x</sub> (bottom) and Pt/WO<sub>x</sub> (top) interfaces, as shown in the inset. The changes in  $\Phi_B^1$  and  $\Phi_B^2$  were estimated from the slopes of  $\ln(I/I^2)$  according to  $1/k_B T$  for the  $I_{ON}$  and  $I_{OFF}$  states of the Pt/WO<sub>x</sub>/W memristor, respectively. **c** Changes

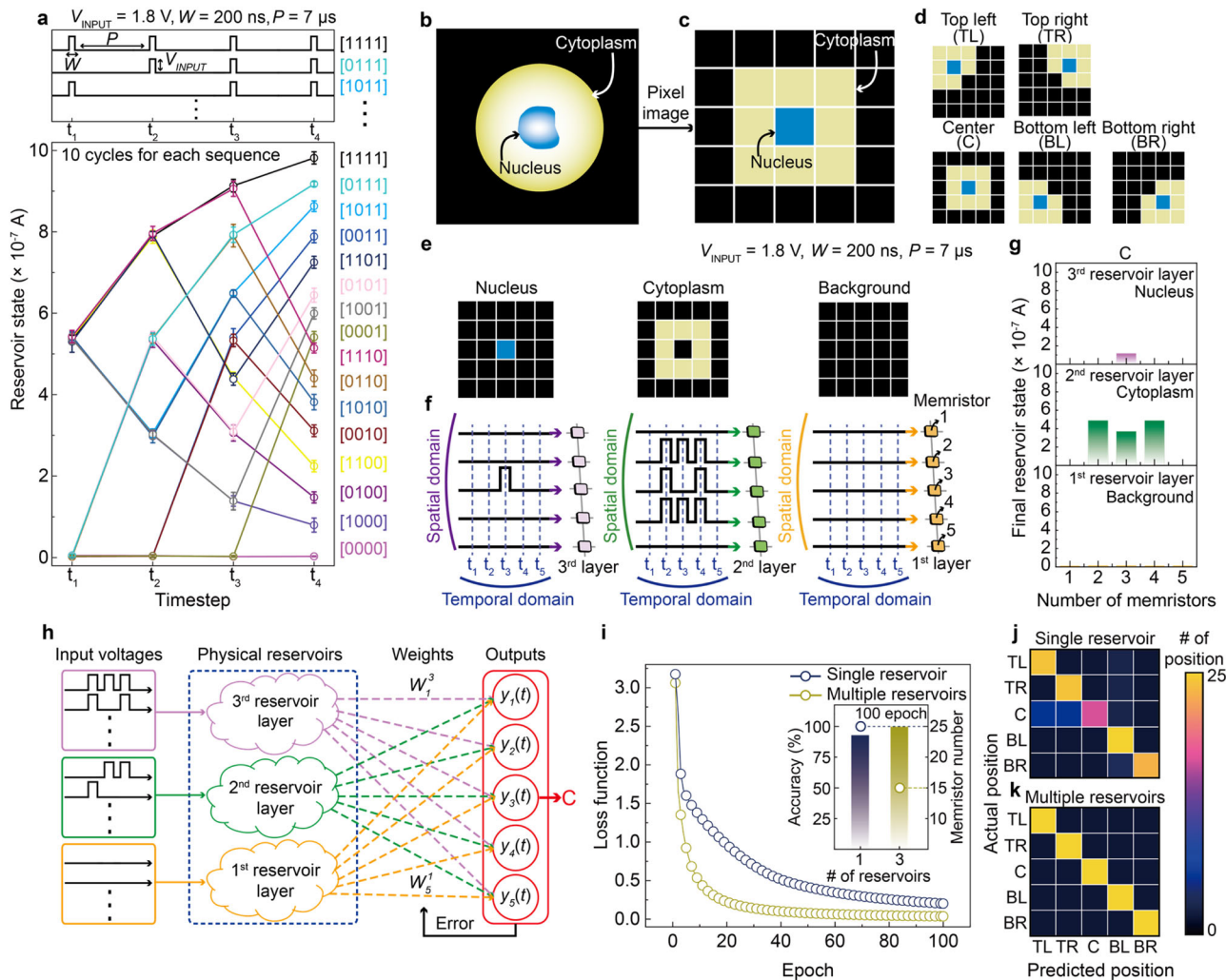
in estimated  $\Phi_B^1$  and  $\Phi_B^2$  for  $I_{ON}$  and  $I_{OFF}$  states. The inset shows the different band diagrams for the  $I_{ON}$  and  $I_{OFF}$  states and state-dependent  $C_{V_o}$ . **d, e** Conductive atomic force microscopy (c-AFM) images of current distributions in  $1 \times 1 \mu\text{m}^2$  WO<sub>x</sub>/W for tip bias ( $V_{bias}$ ) = 3 V (**d**) and  $V_{bias} = -3$  V (**e**). Note that spatial current distribution ( $\bar{I}$ ) was found to be 2.63 nA (**d**) and 8.45 pA (**e**) on the average. The scale bar for both (**d, e**) is 100 nm. **f** Statistical  $I_{ON}$  and  $I_{OFF}$  values for 15 WO<sub>x</sub> memristors according to memristor area.

voltage pulses are applied. Hence, 16 reservoir states are successfully programmed at  $t_4$  without notable cycle variations (10 cycles per sequence, Supplementary Fig. 27). These distinct reservoir states may be attributed to the self-rectifying properties and the consistent cycling switching. The trajectories of change in the reservoir states by the voltage pulse sequences are similarly observed on the memristors from the first to third layers of 3D stacked WO<sub>x</sub> PRs (Supplementary Fig. 28).

To demonstrate the capability of 3D stacked WO<sub>x</sub> PRs to process multiple time series without mutual interference, a detection task of moving biological cells was conducted based on their layer separation in nucleus, cytoplasm, and background, as shown in Figs. 4b–k<sup>56</sup>. Figs. 4b, c show a centered cell image (Fig. 4b) and its simplified  $5 \times 5$ -pixel image (Fig. 4c), where the blue, yellow, and black pixels represent the nucleus, cytoplasm, and background, respectively. As shown in Fig. 4d, the moving cell can be localized in one of five possible cell positions with respect to its nucleus: top left (TL), top right (TR), center (C), bottom left (BL), or bottom right (BR). In three layers of  $5 \times 5$ -pixel images corresponding to the nucleus, cytoplasm, and background images, each layer is encoded according to the five position cells using different voltage sequences in the spatial and time domains. For the centered cell (C) in Figs. 4e, f, various voltage sequences for the five WO<sub>x</sub> memristors in each layer are encoded. The multiple 3D stacked WO<sub>x</sub> PRs can individually capture three local features of the biological cell from the first to third layers. Figure 4g shows the final reservoir states at the center position in terms of these three features. The features are spatiotemporally distinguishable, demonstrating the success of feature extraction using dynamic cell positioning. The results for the other cell positions (TL, TR, BL, and BR) are shown in Supplementary Fig. 29.

When each spatiotemporal input for a cell position enters each reservoir layer, the reservoir state,  $x_n^i$ , in each layer ( $i = 1, 2, 3$ ) is generated. It is then connected to a digital output layer ( $y_1(t), y_2(t), y_3(t), y_4(t)$ , and  $y_5(t)$ ) to perform learning based on the simple single-layer gradient descent (Fig. 4h). Details about learning and prediction are provided in the Methods section. Figure 4i shows the loss function over epoch for a single reservoir and multiple reservoirs. As shown in Fig. 4i, the loss function for multiple reservoirs rapidly decreases with the number of epochs (~4.5 times faster), leading to a higher classification accuracy. Note that the further discussion on the convergence speed is described in Supplementary Note 6. In this detection task, the accuracy of the multiple 3D stacked WO<sub>x</sub> PRs is 100% at 100 epochs, being 7% higher than that of a single reservoir. Thus, for multiple PRs, only the diagonal boxes are fully saturated with yellow in the confusion matrices for the classification results between the actual and predicted positions (Fig. 4k). However, for a single reservoir, the blue saturation of several off-diagonal boxes is observed, indicating incorrect predictions (Fig. 4j). In addition, a single reservoir requires a larger number of memristors (25) than multiple reservoirs (15) for this detection task, and it cannot distinguish between the nucleus and cytoplasm at the overlapping position (Supplementary Fig. 30). Therefore, multiple PRs can extract rich features from time-series data because each reservoir learns distinct local features, even when using 60% fewer memristors than a single reservoir, thus surpassing the performance of a simple increase in packing density. Additional discussions on the wide RC are described in Supplementary Note 2, 6.

As another example of physical RC, we considered a time-dependent Lorenz attractor, which is a nonlinear and deterministic 3D chaotic dynamic system for representing convection in the atmosphere that is difficult to predict<sup>57</sup>. We attempted the Lorenz attractor



**Fig. 4 | Physical reservoir computing for cell position classification.** **a** Change in physical reservoir (PR) states of memristor cell at read voltage ( $V_{\text{READ}} = 0.5 \text{ V}$ ) per timestep ( $t_1 - t_4$ ) and 16 voltage pulse sequences ([1111], [0111], ..., [0000]). A voltage pulse ( $V_{\text{INPUT}}$ ) is 1.8 V for 200 ns with interval ( $P$ ) = 7  $\mu\text{s}$ . Each reservoir state at  $t_4$  is clearly separated regardless of the cycle variation. **b, c** Illustrations of biological cell showing nucleus, cytoplasm, and background (**b**), and its corresponding 5  $\times$  5-pixel image (**c**). **d** Schematics of five cell positions: top left (TL), top right (TR), center (C), bottom left (BL), and bottom right (BR). **e, f** Schematics of three 5  $\times$  5-pixel images of nucleus, cytoplasm, and background for case C (**e**), and their corresponding time-series sequences applied to three PR layers (nucleus for third layer, cytoplasm for second layer, and background for first layer) (**f**). The time-series inputs were

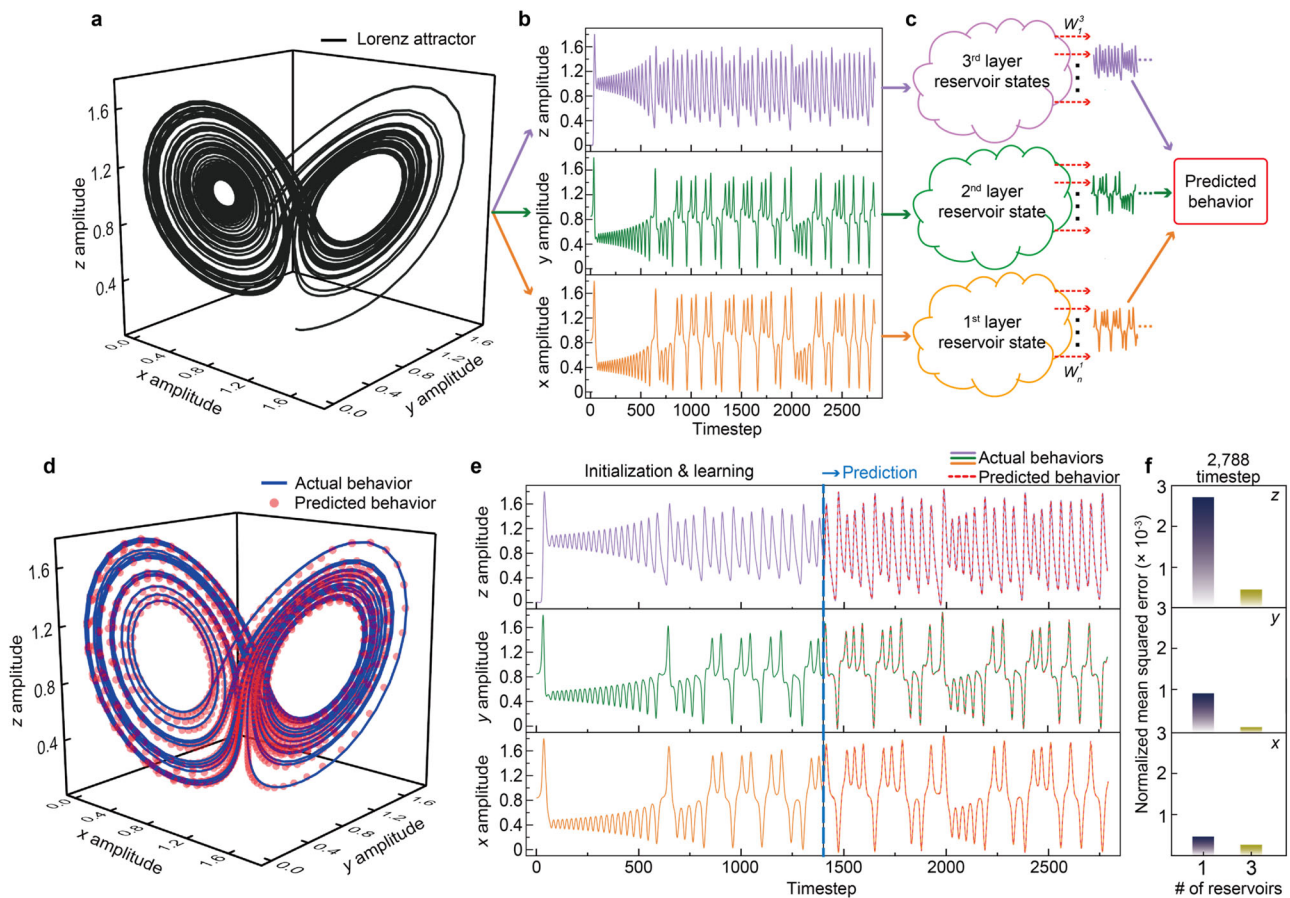
voltage pulses converted from each local feature and introduced into different locations of the memristors in each layer (5 memristors per layer for the spatial domain) at different times ( $t_1 - t_5$  in the temporal domain). **g** Final reservoir states at  $t_5$  obtained from first to third reservoir layers (5 memristors per layer). **h** Schematic of reservoir computing (RC) configuration consisting of input layer for time-series data, three PRs, and output layer. Weights ( $W_i^j$ ) between PRs and output layer was trained to reduce the error (i.e., loss function) by comparing outputs ( $y_n(t)$ ) and the desired value. **i** Loss functions for single (navy) and multiple (yellow) reservoirs according to training epoch. The inset shows the classification accuracies at 100 epochs and numbers of used memristors for the single and multiple reservoirs. **j, k** Confusion matrices for single (**j**) and multiple (**k**) reservoirs.

prediction and simulation using multiple 3D stacked  $\text{WO}_x$  PRs as a proof of concept. Figure 5a shows the 3D trajectory obtained from the Lorenz attractor equations with specific parameters (see the Methods section). As shown in Figs. 5b, c, by deconvoluting the 3D Lorenz equation into three time-dependent one-dimensional equations (i.e.,  $x$ -,  $y$ -, and  $z$ -axis components over a timestep) (Fig. 5b), the 3D stacked  $\text{WO}_x$  PRs could directly handle their corresponding chaotic inputs to generate separable reservoir states, which were then transferred to the output layer (Fig. 5c).

Based on the linear regression algorithm, the learning process was performed at the output layers considering 100 nodes per reservoir (300 nodes in total). As shown in Fig. 5d, after initialization and learning over 1400 timesteps, the predicted behavior closely matches the actual behavior with a negligible deviation. Figure 5e shows the amplitudes of the  $x$ ,  $y$ , and  $z$  components over time for the actual and predicted behaviors (2788 timesteps). Hence, we demonstrate the

successful prediction of a time-dependent 3D Lorenz attractor using multiple 3D stacked  $3 \times 10 \times 10$   $\text{WO}_x$  PRs. The average normalized mean squared error (NMSE) between the predicted and actual behaviors is approximately  $2.62 \times 10^{-4}$ , being >10 times better than that when using a single reservoir with 300 nodes (NMSE of  $1.35 \times 10^{-3}$ ) (Fig. 5f and Supplementary Fig. 31). If the reservoir capacity is insufficient, even when considering multiple reservoirs, the predicted behavior substantially deviates from the actual behavior (Supplementary Fig. 32). Details about prediction are provided in the Methods section. The results indicate that the optimized  $W_n^i$  in the output layer obtained from learning allows to suitably predict the future behavior in response to multiple reservoir states (three layers) at the current timestep. As a result, 3D stacked  $\text{WO}_x$  memristor arrays promote the prediction accuracy and efficiency of physical RC for multiple time-series processing. Note that, in Supplementary Note 6 and Supplementary Figs. 33, 34, we show relevant discussions of physical RC





**Fig. 5 | Physical RC for prediction of time-dependent Lorenz attractor.** **a** 3D trajectory of Lorenz attractor in the  $x$ ,  $y$ , and  $z$  components. **b**, **c** One-dimensional time-dependent Lorenz attractors over time deconvoluted from 3D Lorenz attractor (**b**), and schematic of three physical reservoirs (PRs) for prediction of the 3D Lorenz attractor (**c**). **d** Actual and predicted behaviors of 3D Lorenz attractor after 1400 timesteps. **e** Individual actual behaviors of  $x$ ,  $y$ , and  $z$  components over

time (colored solid lines), and their corresponding prediction results (red dotted lines) using multiple 3D stacked  $3 \times 10 \times 10$   $WO_x$  PRs after 1400 timesteps (blue dotted line). **f** Normalized mean squared error (NMSE) between actual and predicted behaviors at timestep 2788 for single (navy) and multiple (yellow) reservoirs in the  $x$ ,  $y$ , and  $z$  components.

based on multiple 3D stacked  $WO_x$  PRs as compared with traditional approaches and suggest how they could be used at the circuit level.

## Discussion

We design and fabricate a 3D stacked three-layer  $3 \times 10 \times 10$   $WO_x$  memristive crossbar array to implement a wide physical reservoir system for processing multiple time-series information. This design exhibits a forming-free feature, high switching uniformity and device yield (>98%) for a three-layer stacked array, robust endurance (> $10^5$  cycles), and low programming voltage (-0.7 V) for self-rectifying switching. Moreover, it demonstrates essential reservoir properties, including separation and memory fading. The 3D stacked  $WO_x$  memristive crossbar array for wide physical RC allows each array layer to individually portray and process distinct local features of spatio-temporal inputs such that the reservoir states can be readily analyzed at the output layer for classification and prediction tasks. By leveraging this strategy, for dimensional classification of biological cell positions, the multiple PRs achieved better classification accuracy (~100%) and ~4.5 times faster convergence using 60% fewer memristors (15) compared with a single reservoir (93% accuracy with 25 memristors). In addition, for the prediction of the chaotic Lorenz attractor, they achieved a tenfold lower prediction NMSE ( $\sim 2.62 \times 10^{-4}$ ) than a single reservoir (NMSE of  $1.35 \times 10^{-3}$ ). Although the array size is still small as a proof-of-principle demo, both the principles and device structure are applicable to large scales. The proposed 3D stacked  $WO_x$  PR array

represents a leap forward toward an efficient wide RC platform, promising for time-series processing in edge computing.

## Methods

### Fabrication of 3D stacked three-layer memristive crossbar array

The 3D stacked three-layer  $3 \times 10 \times 10$   $WO_x$  memristive crossbar array was fabricated on either a glass substrate or 285 nm-thick  $SiO_2/p^{++}$ -Si substrate ( $1.5 \text{ cm} \times 1.5 \text{ cm}$ ). First, the substrates were prepared using a conventional cleaning process with acetone, isopropyl alcohol, and deionized water via ultra-sonication for 3 min each. To form 1st bottom electrode lines (< 10 nm thick) on the substrate, W was deposited by sputtering system with 5 min pre-sputtering step, gas flow of Ar 10 sccm, working pressure of 10 mTorr, and gun power of 100 W with the patterned shadow mask ( $1.5 \text{ cm} \times 1.5 \text{ cm}$ ) with line width of 100  $\mu\text{m}$ . Note that all the following fabrication processes were performed through sputtering and shadow masks to circumvent the sidewall issue that is critical for the 3D stacked structure. The same-sized shadow masks with the substrate were used to facilitate the patterning alignments. Subsequently, the  $WO_x$  (~180 nm thick) was deposited on the center of the 1st bottom W electrodes except for the contact pads. The patterning alignment was carefully carried out with an optical microscopy. During  $WO_x$  deposition, various combinational sputtering parameters were used, such as 20 min pre-sputtering step, gas flow of Ar 60 sccm and  $O_2$  1 sccm, working pressure of 20 mTorr, gun power of 150 W, and substrate temperature of 40  $^\circ\text{C}$ , to optimize a nonlinear

dynamic switching with self-rectification (see Supplementary Fig. 17). Perpendicularly to the 1st bottom W electrode lines, the 1st top Pt electrode lines (<10 nm) were formed on the 1st WO<sub>x</sub> layer by using a shadow mask and sputter. The Pt was deposited with 5 min pre-sputtering step, gas flow of Ar 10 sccm, working pressure of 10 mTorr, and gun power of 100 W. After the Pt deposition, W was directly sputtered on the 1st top Pt electrodes to form 2nd bottom W electrode lines. The 2nd WO<sub>x</sub> layer was then formed on the 2nd bottom W electrode/1st top Pt electrode lines, except for the all electrode pads. The same sputtering parameters used during the formation of 1st WO<sub>x</sub> were employed. Similar fabrication steps were repeated until completing the 3D stacked of the three-layer memristive crossbar array. The top-view optical images of the fabrication processes and corresponding schematics are shown in Supplementary Fig. 3.

### Structural analysis

To identify the formation of 3D stacked junction structure at a certain crosspoint, the sample for TEM and EDS analysis was prepared via focused ion beam technique (FIB, Helios G4 HX) and then cross-sectional TEM image and EDS results of the 3D stacked memristors were obtained by using a FEI Double Cs and monochromated TEM instrument. The depth-profiling XPS analysis (PHI 5000 VersaProbe) was performed by using 100 μm-diameter sized single Pt/WO<sub>x</sub>/W memristor (80-nm thick WO<sub>x</sub>).

### Electrical characterization

A semiconductor parameter analyzer (4155 C, Agilent), pulse generator (81104 A, Keysight), and low-leakage switch mainframe (E5250A, Keysight) were used to electrically investigate and characterize the *I*-*V* switching curves, consecutive sweeps, endurance cycles, voltage ranges, temperature-dependent Schottky barrier height, and size dependency of the 3D stacked three-layer 3 × 10 × 10 WO<sub>x</sub> memristive crossbar array. Memristors in the 3D stacked arrays were randomly accessible when contacting their corresponding electrode pads. For example, to select a certain memristor of the 1st layer, the 1st bottom W electrode pad was grounded, whereas the 1st top Pt & 2nd bottom W electrode pad was biased. Considering that the 1st top Pt electrode is underneath the 2nd bottom W electrode, the Pt/WO<sub>x</sub>/W memristor of the 1st layer can be characterized. This measurement can be considered a reasonable method because our WO<sub>x</sub>-based memristors operate based on interfacial switching. The time-decay features and current responses (i.e., reservoir states) as functions of various time-dependent inputs were measured by using a Keithley 4200 semiconductor characterization system that can program arbitrary voltage schemes into memristors. To investigate self-decaying nature, a *V*<sub>READ</sub> of 0.5 V for 5 μs was first applied into the selected memristor in the 3D stacked array to confirm *I*<sub>OFF</sub> state. Then, a programming pulse with amplitude of 1.8 V for pulse widths (*W*) = 1 ms, 100 μs or 50 μs, which can switch the device into *I*<sub>ON</sub> state, was introduced into the memristor. The rising and falling edges of the pulses were set to be 500 ns. After the programming, a *V*<sub>READ</sub> of 0.5 V was used to investigate the current decaying behaviors over time. To generate reservoir states, time-varying voltage pulse schemes consisting of four timesteps were used (Fig. 4a). As shown in the upper panel of Fig. 4a, 16 binary encoding sequences (1 or 0) consist of 4 timesteps (t<sub>1</sub>-t<sub>4</sub>) at *P* = 7 μs, and each timestep contains either a *V*<sub>INPUT</sub> of 1.8 V for *W* = 200 ns when its corresponding encoding is 1 or no pulse when its corresponding encoding is 0. The rising and falling edges of the programming pulse were set to be 50 ns. After each timestep, *V*<sub>READ</sub> of 0.5 V at the time interval of 5 ns was used to record reservoir states (i.e., current behaviors). Note that for the random access, the same method used in the electrical characterization was employed.

### Conductive atomic force microscopy

To identify the filament-free and homogeneous switching phenomena of the WO<sub>x</sub> memristor, conductive atomic force microscopy (c-AFM)

measurement was performed. In order to map the spatial current distribution ( $\bar{I}$ ), a Pt/Ir-coated c-AFM tip was used to directly scan the top surface of the WO<sub>x</sub>/W sample and configure the Pt/WO<sub>x</sub>/W junction structure (Supplementary Fig. 25). The Pt/Ir-coated tip was set to be grounded, and the bottom W electrode was set to be biased; hence, tip bias (*V*<sub>bias</sub>) was represented as the opposite bias polarity. The tip (-24 nm) and estimated contact radii (-2.1 nm) are discussed in the Supplementary Information (Supplementary Fig. 25). The electrical characteristics were measured in scanning mode with a scan area of 1 × 1 μm<sup>2</sup> by using a DLPCA-2000 built-in current amplifier (Electro-Optical Components) under a humidity below 15%. After the Pt/Ir tip contacted the WO<sub>x</sub>/W samples, -3 V and 3 V were applied to the bottom W electrode during scanning, respectively. Then, the spatial current distribution ( $\bar{I}$ ) at *V*<sub>bias</sub> = 3 V and -3 V were mapped on the 1 × 1 μm<sup>2</sup> scan area and it was verified that there is no occurrence of a conductive filament.

### Estimation of Schottky barrier height

To determine the Schottky barrier heights of the Pt/WO<sub>x</sub>/W memristor, temperature-dependent *I*-*V* curves were investigated and plotted according to the following thermionic emission equations:

$$I = AA^* T^2 \exp\left(\frac{-q\Phi_B}{K_B T}\right), \quad (1)$$

$$\ln \frac{I}{T^2} = \ln(AA^*) - \frac{q\Phi_B}{K_B T}, \quad (2)$$

where *A*, *A*<sup>\*</sup>, *T*, *q*,  $\Phi_B$ , and *K<sub>B</sub>* are the effective area, Richardson constant, temperature, electronic charge, Schottky barrier height, and Boltzmann constant, respectively. Hence,  $\Phi_b^1$  and  $\Phi_b^2$  for the ON and OFF states at *V*<sub>READ</sub> = ± 0.5 V are quantitatively estimated from the slope of the straight fitting lines shown in Fig. 3b.

### Dimensional classification of biological cell positions

The dimensional classification of five biological cell positions was simulated based on the individual PR states encoded from each layer of the 3D stacked WO<sub>x</sub> array. Three layers of 5 × 5-pixel images corresponding to the nucleus, cytoplasm, and background images are individually encoded by each layer of the 3D stacked WO<sub>x</sub> array. A Keithley 4200 system, capable of generating arbitrary voltage pulse patterns, was employed to measure the reservoir states (i.e., current responses) at each timestep (from t<sub>1</sub> to t<sub>5</sub>) by applying the spatio-temporal inputs shown in Figs. 4e-g. Five memristors per layer were selected in each 10 × 10 crossbar array for spatial information processing of the biological cell, and all reservoir states were read at *V*<sub>READ</sub> past 5 μs after the application of each programming pulse (1.8 V for 200 ns). The programming and reading disturbances in the 3D stacked 3 × 10 × 10 WO<sub>x</sub> array can be minimized owing to the self-rectifying function. The reservoir states of the different memristors in each layer were individually recorded by applying time-dependent voltage pulses to the top Pt electrode line of the selected memristor with a grounded bottom W electrode line.

Subsequently, a single-layer neural network with a supervised learning algorithm was employed as the output layer. By considering the PR states at the final time frame (t<sub>5</sub>) as input *x*, output *y* was generated through the softmax activation function:

$$y_j = \frac{e^{z_j}}{\sum_{k=1}^j e^{z_k}}, z_j = W_{ij} \cdot x_i + b, \quad (3)$$

where *W*<sub>*ij*</sub> is the synaptic weight between a reservoir and output layer, *b* is the bias, and *i* and *j* are integers identifying the reservoir nodes and output neurons, respectively. The number of output neurons was five, corresponding to the considered cell positions (top left (TL), top right



(TR), center (C), bottom left (BL), or bottom right (BR)). During learning, the cross-entropy loss function ( $\varepsilon$ ) was decreased using delta rule of gradient descent as follows:

$$\varepsilon = -\frac{1}{n} \sum_{m=1}^n \sum_{k=1}^j \hat{y}_k^{(m)} \log(y_k^{(m)}), \quad (4)$$

$$W_{ij} \equiv W_{ij} - \eta \frac{\partial \varepsilon}{\partial W}, \quad (5)$$

where  $n$  is the number of input data points,  $\hat{y}_j$  is the one-hot encoding of  $y_j$ , and  $\eta$  is the learning rate. After learning, additional images were generated per cell position and used during prediction (Supplementary Fig. 35), and also their corresponding PR states were produced based on the fitting results (Supplementary Fig. 36).

### Prediction of time-dependent Lorenz attractor

The 3D Lorenz attractor is a set of chaotic solutions of the Lorenz system described by three ordinary differential equations according to the axis:

$$\frac{dx}{dt} = \alpha(y - x), \quad (6)$$

$$\frac{dy}{dt} = x(\beta - z) - y, \quad (7)$$

$$\frac{dz}{dt} = xy - \gamma z. \quad (8)$$

These equations can describe various chaotic behaviors and physical phenomena over time depending on the values of parameters  $\alpha$ ,  $\beta$ , and  $\gamma$ . By setting  $\alpha = 10$ ,  $\beta = 28$ , and  $\gamma = 8/3$  for the chaotic behaviors, the Lorenz attractor was normalized and deconvoluted into time-varying amplitudes for the  $x$ ,  $y$ , and  $z$  components. These values were applied into each PR after normalizing the inputs in the operating voltage range from 0 V to 1.8 V, as shown in Figs. 5b, c. At each reservoir, 100 memristors were utilized and simulated to enrich the reservoir states because some device variations can produce qualitatively similar but quantitatively different current responses (Supplementary Fig. 28 and Supplementary Note 2, 6). The reservoir states obtained from the three PRs were fed into the corresponding output layers, and the weights were optimized using the supervised learning algorithm of linear regression with the pseudo inverse:

$$W = (X^T X)^{-1} X^T Y \quad (9)$$

$X$ ,  $W$ , and  $Y$  are a set of reservoir states, weights, and actual behaviors, respectively. After initialization during the initial 100 timesteps, the weights were updated until 1400 timesteps. Initialization was required to prepare and assist the prediction processing of the RC system without learning because the initial conditions of the Lorenz attractor could strongly affect its future behavior. When learning was complete, the actual behaviors at the current timestep ( $t$ ) were used to predict the behaviors at the next timestep ( $t + 1$ ).

### Normalized mean squared error

The normalized mean squared error (NMSE) used for performance evaluation was calculated as

$$\text{NMSE} = \frac{\sum_t \sum_j (y_j(t) - \hat{y}_j(t))^2}{\sum_t \sum_j \hat{y}_j^2(t)}, \quad (10)$$

where  $t_j(t)$  and  $y_j(t)$  are the actual and predicted behaviors, respectively.

### Data availability

All data supplementary to the findings of this study are available within the article and its Supplementary Information, or from the corresponding author upon request.

### Code availability

All codes used for the simulation are available within the article and its Supplementary Information, or from the corresponding author upon request.

### References

1. Analytics, M. The age of analytics: competing in a data-driven world. *McKinsey Global Institute Research* (2016).
2. Reichstein, M. et al. Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).
3. Escalona-Morán, M. A., Soriano, M. C., Fischer, I. & Mirasso, C. R. Electrocardiogram classification using reservoir computing with logistic regression. *IEEE J. Biomed. Health Inform.* **19**, 892–898 (2014).
4. Mammedov, Y. D., Olugu, E. U. & Farah, G. A. Weather forecasting based on data-driven and physics-informed reservoir computing models. *Environ. Sci. Pollut. Res.* 1–14 (2021).
5. Ilies, I. et al. Stepping forward through echoes of the past: forecasting with echo state networks. <http://www.neural-forecasting-competition.com/downloads/NN3/methods/27-NN3HerbertJaegerreport.pdf> (Jacobs University Bremen, 2007).
6. Sacchi, R., Ozturk, M. C., Principe, J. C., Carneiro, A. A. & da Silva, I. N. Water Inflow Forecasting using the Echo State Network: a Brazilian Case Study, In *2007 International Joint Conference on Neural Networks (IJCNN)* 2403–2408 (IEEE, 2007).
7. Newbold, P. ARIMA model building and the time series analysis approach to forecasting. *J. Forecast.* **2**, 23–35 (1983).
8. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
9. Tanaka, G. et al. Recent advances in physical reservoir computing: a review. *Neural Netw.* **115**, 100–123 (2019).
10. Nakajima, K. Physical reservoir computing—an introductory perspective. *Jpn J. Appl. Phys.* **59**, 060501 (2020).
11. Kulkarni, M. S. & Teuscher, C. Memristor-based Reservoir Computing In *2012 IEEE/ACM International Symposium on Nanoscale Architectures*. (NANOARCH) 226–232 (2012).
12. Du, C. et al. Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **8**, 2204 (2017).
13. Midya, R. et al. Reservoir computing using diffusive memristors. *Adv. Intell. Syst.* **1**, 1900084 (2019).
14. Moon, J. et al. Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat. Electron.* **2**, 480–487 (2019).
15. Sun, L. et al. In-sensor reservoir computing for language learning via two-dimensional memristors. *Sci. adv.* **7**, eabg1455 (2021).
16. Zhong, Y. et al. Dynamic memristor-based reservoir computing for high-efficiency temporal signal processing. *Nat. Commun.* **12**, 408 (2021).
17. Milano, G. et al. In materia reservoir computing with a fully memristive architecture based on self-organizing nanowire networks. *Nat. Mater.* **21**, 195–202 (2022).
18. Zhong, Y. et al. A memristor-based analogue reservoir computing system for real-time and power-efficient signal processing. *Nat. Electron.* **5**, 672–681 (2022).
19. Liu, K. et al. Multilayer Reservoir Computing Based on Ferroelectric  $\alpha$ -In<sub>2</sub>Se<sub>3</sub> for Hierarchical Information Processing. *Adv. Mater.* **34**, 2108826 (2022).

20. Sun, W. et al. 3D Reservoir Computing with High Area Efficiency (5.12 TOPS/mm<sup>2</sup>) Implemented by 3D Dynamic Memristor Array for Temporal Signal Processing. In *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*. 222–223 (IEEE).
21. Chen, A. & Lin, M.-R. Variability of resistive switching memories and its impact on crossbar array performance. In *2011 International Reliability Physics Symposium*. MY. 7.1-MY. 7.4 (IEEE, 2011).
22. Yu, S., Guan, X. & Wong, H.-S. P. On the switching parameter variation of metal oxide RRAM—Part II: Model corroboration and device design strategy. *IEEE Trans. Electron Devices* **59**, 1183–1188 (2012).
23. Xia, Q. & Yang, J. J. Memristive crossbar arrays for brain-inspired computing. *Nat. Mater.* **18**, 309–323 (2019).
24. Choi, S., Yang, J. & Wang, G. Emerging memristive artificial synapses and neurons for energy-efficient neuromorphic computing. *Adv. Mater.* **32**, 2004659 (2020).
25. Han, M. & Mu, D. Multi-reservoir Echo State Network with Sparse Bayesian Learning. In *Advances in Neural Networks-ISNN 2010: 7th International Symposium on Neural Networks, ISNN 2010, Shanghai, China, June 6-9, 2010, Proceedings, Part I* 7. 450–456 (Springer, 2010).
26. Meftah, B., Lézoray, O. & Benyettou, A. Novel approach using echo state networks for microscopic cellular image segmentation. *Cogn. Comput.* **8**, 237–245 (2016).
27. Carmichael, Z., Syed, H., Burtner, S. & Kudithipudi, D. Mod-DeepESN: Modular Deep Echo State Network. In *2018 Conference on Cognitive Computational Neuroscience*. (2018).
28. Yang, Y. & Huang, R. Probing memristive switching in nanoionic devices. *Nat. Electron.* **1**, 274–287 (2018).
29. Di Martino, G. et al. Real-time in situ optical tracking of oxygen vacancy migration in memristors. *Nat. Electron.* **3**, 687–693 (2020).
30. Wang, G. et al. Three-dimensional networked nanoporous Ta<sub>2</sub>O<sub>5-x</sub> memory system for ultrahigh density storage. *Nano Lett.* **15**, 6009–6014 (2015).
31. Choi, S. et al. A self-rectifying TaO<sub>y</sub>/nanoporous TaO<sub>x</sub> memristor synaptic array for learning and energy-efficient neuromorphic systems. *NPG Asia Mater.* **10**, 1097–1106 (2018).
32. Sun, B. et al. Non-zero-crossing current-voltage hysteresis behavior in memristive system. *Mater. Today Adv.* **6**, 100056 (2020).
33. Kim, K. M. et al. Low-power, self-rectifying, and forming-free memristor with an asymmetric programming voltage for a high-density crossbar application. *Nano Lett.* **16**, 6724–6732 (2016).
34. Wu, C., Kim, T. W., Choi, H. Y., Strukov, D. B. & Yang, J. J. Flexible three-dimensional artificial synapse networks with correlated learning and trainable memory capability. *Nat. Commun.* **8**, 752 (2017).
35. Luo, Q. et al. Self-rectifying and forming-free resistive-switching device for embedded memory application. *IEEE Electron Device Lett.* **39**, 664–667 (2018).
36. Li, C. et al. Three-dimensional crossbar arrays of self-rectifying Si/SiO<sub>2</sub>/Si memristors. *Nat. Commun.* **8**, 15666 (2017).
37. Wang, W. et al. A self-rectification and quasi-linear analogue memristor for artificial neural networks. *IEEE Electron Device Lett.* **40**, 1407–1410 (2019).
38. Kim, Y. et al. Kernel application of the stacked crossbar array composed of self-rectifying resistive switching memory for convolutional neural networks. *Adv. Intell. Syst.* **2**, 1900116 (2020).
39. Wu, C. et al. Self-rectifying resistance switching memory based on a dynamic p-n junction. *Nanotechnology* **32**, 085203 (2020).
40. Choi, S. et al. Low-power self-rectifying memristive artificial neural network for near internet-of-things sensor computing. *Adv. Electron. Mater.* **7**, 2100050 (2021).
41. Huang, J.-N., Huang, H.-M., Xiao, Y., Wang, T. & Guo, X. Memristive devices based on Cu-doped NbO<sub>x</sub> films with large self-rectifying ratio. *Solid State Ion.* **369**, 115732 (2021).
42. Jeon, K. et al. Self-rectifying resistive memory in passive crossbar arrays. *Nat. Commun.* **12**, 2968 (2021).
43. Cheong, W. H. et al. Demonstration of neuromodulation-inspired stashing system for energy-efficient learning of spiking neural network using a self-rectifying memristor array. *Adv. Funct. Mater.* **32**, 2200337 (2022).
44. Kim, S. E. et al. Sodium-doped titania self-rectifying memristors for crossbar array neuromorphic architectures. *Adv. Mater.* **34**, 2106913 (2022).
45. Park, S.-O., Jeong, H., Park, J., Bae, J. & Choi, S. Experimental demonstration of highly reliable dynamic memristor for artificial neuron and neuromorphic computing. *Nat. Commun.* **13**, 2888 (2022).
46. Yao, X. et al. Protonic solid-state electrochemical synapse for physical neural networks. *Nat. Commun.* **11**, 3134 (2020).
47. Biju, K. P. et al. Bipolar resistance switching in the Pt/WO<sub>x</sub>/W non-volatile memory devices. *Curr. Appl. Phys.* **11**, e62–e65 (2011).
48. Biju, K. P. et al. Resistive switching characteristics and mechanism of thermally grown WO<sub>x</sub> thin films. *J. Appl. Phys.* **110**, 064505 (2011).
49. Kim, S. et al. Effect of scaling WO<sub>x</sub>-based RRAMs on their resistive switching characteristics. *IEEE Electron Device Lett.* **32**, 671–673 (2011).
50. Wang, Z. et al. Toward a generalized Bienenstock-Cooper-Munro rule for spatiotemporal learning via triplet-STDP in memristive devices. *Nat. Commun.* **11**, 1510 (2020).
51. Wang, G. et al. Hydrogen-treated WO<sub>3</sub> nanoflakes show enhanced photostability. *Energy Environ. Sci.* **5**, 6180–6187 (2012).
52. Katoh, M. & Takeda, Y. Chemical state analysis of tungsten and tungsten oxides using an electron probe microanalyzer. *Jpn J. Appl. Phys.* **43**, 7292 (2004).
53. Lambert-Mauriat, C., Oison, V., Saadi, L. & Aguir, K. Ab initio study of oxygen point defects on tungsten trioxide surface. *Surf. Sci.* **606**, 40–45 (2012).
54. Waser, R. & Aono, M. Nanoionics-based resistive switching memories. *Nat. Mater.* **6**, 833–840 (2007).
55. Peng, H. et al. Electrode dependence of resistive switching in Mn-doped ZnO: filamentary versus interfacial mechanisms. *Appl. Phys. Lett.* **96**, 192113 (2010).
56. Khamael, A.-D. et al. Segmentation of white blood cell, nucleus and cytoplasm in digital haematology microscope images: A review—challenges, current and future potential techniques. *IEEE Rev. Biomed. Eng.* **14**, 290–306 (2020).
57. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).

## Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF-2022R1A2B5B02001455, NRF-2022M3H4A1A01009526, NRF-2022K2A9A1A01098066, and RS-2023-00220077), the Korea Institute of Science and Technology (KIST) Institutional Program (Project No. 2E32491-23-112), the KU-KIST Research Fund, and a Korea University grant.

## Author contributions

G.W. and S.C. conceived the research and wrote the manuscript using inputs from all the authors. S.C. performed the experiments, analyses, and neural network simulations. J.S. and J.S.E. performed the c-AFM measurements on the WO<sub>x</sub>/W sample. G.P. and J.J. assisted in the execution of some experiments, such as WO<sub>x</sub> deposition and electrical characterization. J.J.Y. participated in the discussions regarding the experimental results and provided an interpretation of them. G.W. oversaw the project, revised the manuscript, and led the efforts toward completion with J.J.Y.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains Supplementary Material available at <https://doi.org/10.1038/s41467-024-46323-7>.

**Correspondence** and requests for materials should be addressed to J. Joshua Yang or Gunuk Wang.

**Peer review information** *Nature Communications* thanks Xiangshui Miao, Linfeng Sun, Ilia Valov and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

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

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

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

© The Author(s) 2024