

Supplementary Materials for  
**Thin-film transistor for temporal self-adaptive reservoir computing with  
closed-loop architecture**

Ruiqi Chen *et al.*

Corresponding author: Peng Huang, [phwang@pku.edu.cn](mailto:phwang@pku.edu.cn); Xiaoyan Liu, [xyliu@ime.pku.edu.cn](mailto:xyliu@ime.pku.edu.cn)

*Sci. Adv.* **10**, eadl1299 (2024)  
DOI: 10.1126/sciadv.adl1299

**This PDF file includes:**

Note S1  
Figs. S1 to S22  
Table S1  
References

### Supplementary Note 1. Physical mechanism of the charge redistribution in the ZnO layer

The alteration of effective gate voltage  $\Delta V$  that results from the positive/negative charge in the interlayer is calculated following Gaussian's law:

$$\Delta V = -\frac{1}{A} \int_{d_1}^{d_2} \frac{1}{\epsilon_{eff}(x)} x \rho(x) dx \quad (1)$$

where  $\epsilon_{eff}(x)$  denotes the effective dielectric constant between the gate and the location of charge, while  $A$  is the area of gate. It can be clearly deduced that the charge located near the channel will induce more prominent alteration in the gate voltage. We comprehensively consider the influence of negative/positive charges that distribute in the interlayer, and write the subsequent effective gate voltage  $V_{Geff}$  as:

$$V_{Geff} = V_{GS0} + \frac{Q}{C_{eff}} \quad (2)$$

in which  $Q$  is the charge located near the channel.  $C_{eff}$  is the effective capacitance.

In the following, we investigate how a single pulse of gate voltage impacts the drain current. It is commonly agreed that when a thin film transistor works in subthreshold region, the drain current is written as:

$$I_d = \frac{W}{L} D_n \frac{kT}{q} C_D \exp \left[ \frac{q(V_{Geff} - V_{th})}{nkT} \right] \left[ 1 - \exp \left( -\frac{qV_D}{kT} \right) \right] \quad (3)$$

where  $n$  stands for the ideality factor. For brevity, the above equation can be written as follows:

$$I_d = I_0 \exp \left[ \frac{qV_{Geff}}{nkT} \right] \quad (4)$$

When applying a single pulse of positive voltage to the gate, the charge distribution in the ZnO interlayer is altered, resulting in an abrupt increment  $\Delta I_d$  in drain current. In the following part, we only discuss voltage pulses with fixed amplitude and width. Denoting the pulse number with  $N$ , we can analyze the magnitude of  $\Delta I_d$  from the perspective of  $Q$ :

$$\begin{aligned}
\Delta I_d &= \frac{\partial I_d}{\partial V_{Geff}} \cdot \frac{\partial V_{Geff}}{\partial Q} \cdot \Delta Q \\
&= I_d \frac{q}{nkT} \frac{1}{C_{eff}} \Delta Q
\end{aligned} \tag{5}$$

$\Delta Q$  is the charge quantity of electrons that are activated away from  $V_o$  defect levels during one single positive voltage pulse, which will naturally be dependent on many factors, such as the properties of input pulse and the amount of remaining  $V_o$  that can be activated. Since the properties of pulse are fixed, we only need to consider the latter factor. In order to analyze  $\Delta Q$ , we consider two extreme cases: (1) No electrons are activated yet. There is abundant  $V_o$ . (2) No more electrons can be activated away from  $V_o$  at current pulse condition.  $V_o$  is depleted.

In case (1), since there exists plenty of  $V_o$  for ionization, the quantity of activated electrons should be constant as  $N$  increases. Therefore, in this situation we have  $\Delta Q = C_I$ , where  $C_I$  represents a constant.

In case (2), however,  $\Delta Q$  will gradually decrease to zero because of saturation. When applying positive pulses, the probability for electrons at energy level  $\varepsilon$  to activate away from  $V_o$  is

$$\begin{aligned}
P(\varepsilon) &\propto \exp\left(-\frac{E_B - \varepsilon}{kT}\right) \\
&= P_0 \exp\left(\frac{\varepsilon}{kT}\right)
\end{aligned} \tag{6}$$

Assuming the density of states for  $V_o$  to be uniform, i.e.,  $N_{V_o}(\varepsilon) = N_{V_o}(E_{min} < \varepsilon < E_{max})$ ,

we can deduce that the change of charge quantity during the time of  $\Delta t$  is:

$$\begin{aligned}
\Delta Q &= q \int_{E_F}^{E_0} P(\varepsilon) \Delta t N_{V_o}(\varepsilon) d\varepsilon \\
&= q P_0 N_{V_o} \Delta t \int_{E_F}^{E_0} \exp\left(\frac{\varepsilon}{kT}\right) d\varepsilon \\
&= q k T P_0 N_{V_o} \Delta t \left[ \exp\left(\frac{E_0}{kT}\right) - \exp\left(\frac{E_F}{kT}\right) \right] \\
&= q k T P_0 N_{V_o} \Delta t \left[ \exp\left(\frac{(Q_m - Q) / q N_{V_o} + E_F}{kT}\right) - \exp\left(\frac{E_F}{kT}\right) \right] \\
&= q k T P_0 N_{V_o} \Delta t \exp\left(\frac{E_F}{kT}\right) \left[ \exp\left(\frac{Q_m - Q}{q N_{V_o} kT}\right) - 1 \right] \\
&= C_2 \exp\left[C_3 (Q_m - Q) - 1\right]
\end{aligned} \tag{7}$$

Where  $E_0$  denotes the current edge between  $V_o$  and ionized  $V_o^+/V_o^{2+}$ .  $Q_m$  denotes the maximum charge quantity that can be obtained under this pulse condition. For simplicity, we write (7) as

$$\Delta Q = C_2 \exp\left[C_3 (Q_m - Q) - 1\right] \tag{8}$$

Finally, we obtain:

$$\Delta Q = \begin{cases} C_1 & Q \rightarrow 0 \\ C_2 \exp\left[C_3 (Q_m - Q) - 1\right] & Q \rightarrow Q_m \end{cases} \tag{9}$$

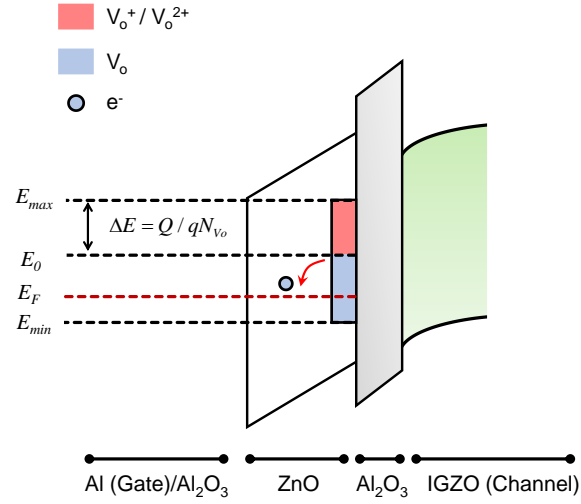
This equation indicates when  $Q \rightarrow 0$ , i.e., effects from voltage pulses have not substantially accumulated and  $I_d$  is close to its primitive value, we have  $\Delta I_d \propto I_d$ . When  $Q \rightarrow Q_m$ , i.e.,  $I_d$  is largely enhanced by gate pulses and is approaching its maximum value,  $\Delta I_d$  will decrease with  $I_d$  and finally be zero. That is when the maximum of  $I_d$  is achieved. In conclusion, we have

$$\Delta I_d = \begin{cases} C_1 I_d & I_d \rightarrow I_{d,min} \\ 0 & I_d \geq I_{d,max} \end{cases} \tag{10}$$

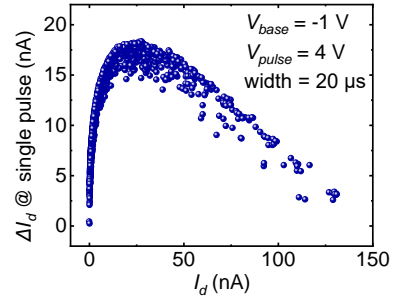
**This conclusion agrees well with the experimental result** (Fig. S2).

Following the analysis of mechanisms, we have looked into the possible directions for enlarging the modulation range of the temporal constant  $\tau$  as follows. The relaxation process is related to the trap emptying and filling process in the amorphous oxide

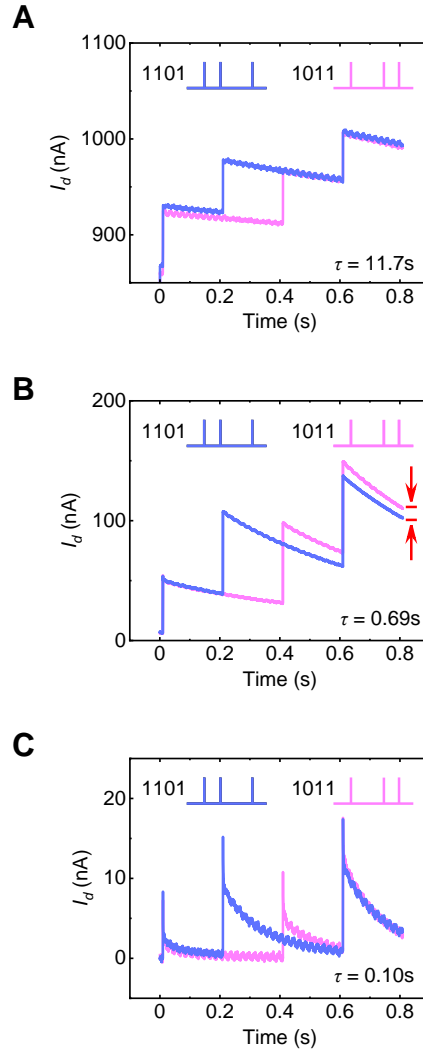
semiconductor channel. According to precedent researches, the effective electronic field is responsible for enhancing the de-trapping probability of electrons accommodated above the Fermi level, and also the trapping probability of that below the Fermi level (55, 56). Consequently, for our device, the relaxation rate is enhanced as the electronic field in ZnO increases. Therefore, in order to extend the modulation range of the relaxation temporal constant, one should aim to enlarge the effect that the gate voltage can have on the electric field in ZnO, or in other words, the proportion of voltage drop on the ZnO layer. Since the gate voltage is mainly distributed to the insulator-ZnO-insulator structure, several optimization solutions that minimize the effective oxide thickness (EOT) can be considered, including decreasing the thickness of  $\text{Al}_2\text{O}_3$  or substituting it with other high-k materials with higher dielectric constant. The trade-off between the gate leakage current needs to be considered for the former method. Meanwhile, the interface quality should be taken care of for the latter. In conclusion, we believe that decreasing the thickness of the insulating material or replacing it with materials with a higher dielectric constant would improve the modulation range of the time constant.



**Fig. S1 Bandgap diagram illustrating the transferring process of electrons during the positive gate voltage pulse.** The total quantity of transferred electrons is determined by time duration and the integration of hopping probability throughout different energy levels.

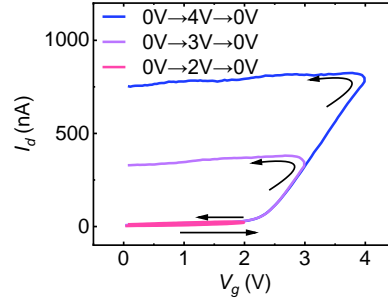


**Fig. S2 Experimental result of the  $\Delta I_d$ - $I_d$  relation.** The experimental data agrees well with theoretical analysis.

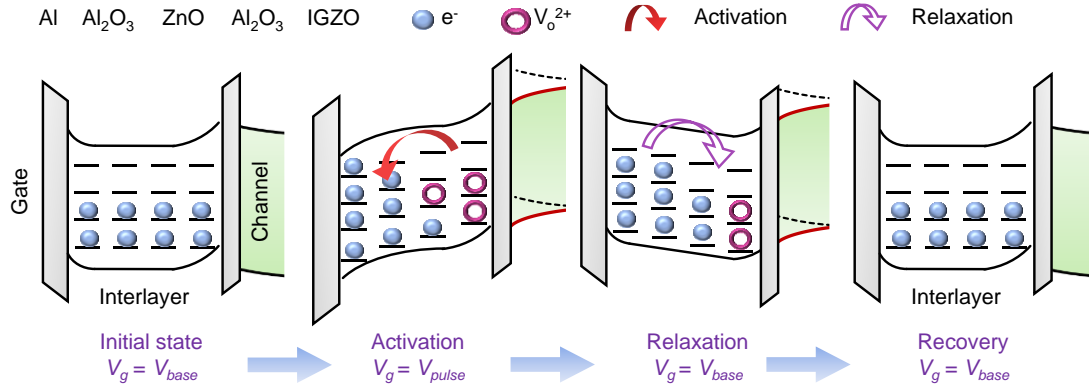


**Fig. S3 Experimental demonstration of how the temporal dynamics of reservoir affect the separability of reservoir states.** Two input sequences “1101” and “1011” are applied to the same FM TFT device. The sequence is inputted in the form of gate voltage spike with a frequency of 5 Hz. In (A)-(C) the same FM TFT device was used, while different  $V_{base}$  was chosen. The three conditions exhibit different timescales with  $\tau_1=11.7$  s @  $V_{base}=1$  V,  $\tau_2=0.69$  s @  $V_{base}=0$  V,  $\tau_3=0.10$  s @  $V_{base}=-0.5$  V. For this specific input, in both the  $\tau_1$  and  $\tau_3$  situations, the final current the two distinct input overlaps. This means that the reservoir states will be indistinguishable if  $\tau$  is either too long or too short. Only with moderate  $\tau$  ( $\tau_2$ ) can the reservoir generate separable reservoir states for the two different inputs ‘1101’ and ‘1011’.

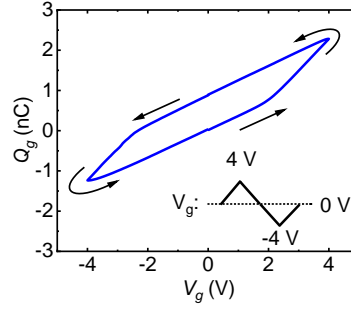




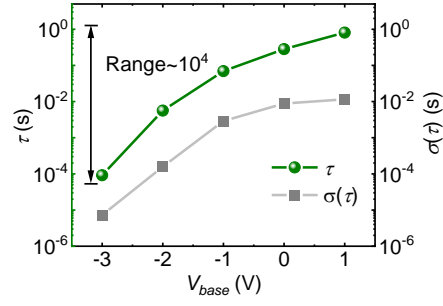
**Fig. S4 Drain current hysteresis in the double sweep of different maximum gate voltage.** The hysteresis of the drain current ( $I_d$ ) grows as the sweeping range of gate voltage ( $V_g$ ) increases. This enables us to control the memory window by adjusting the magnitude of  $V_g$ . The dependence indicates that with a higher magnitude of  $V_g$ , the redistribution of charge in ZnO is more substantial.



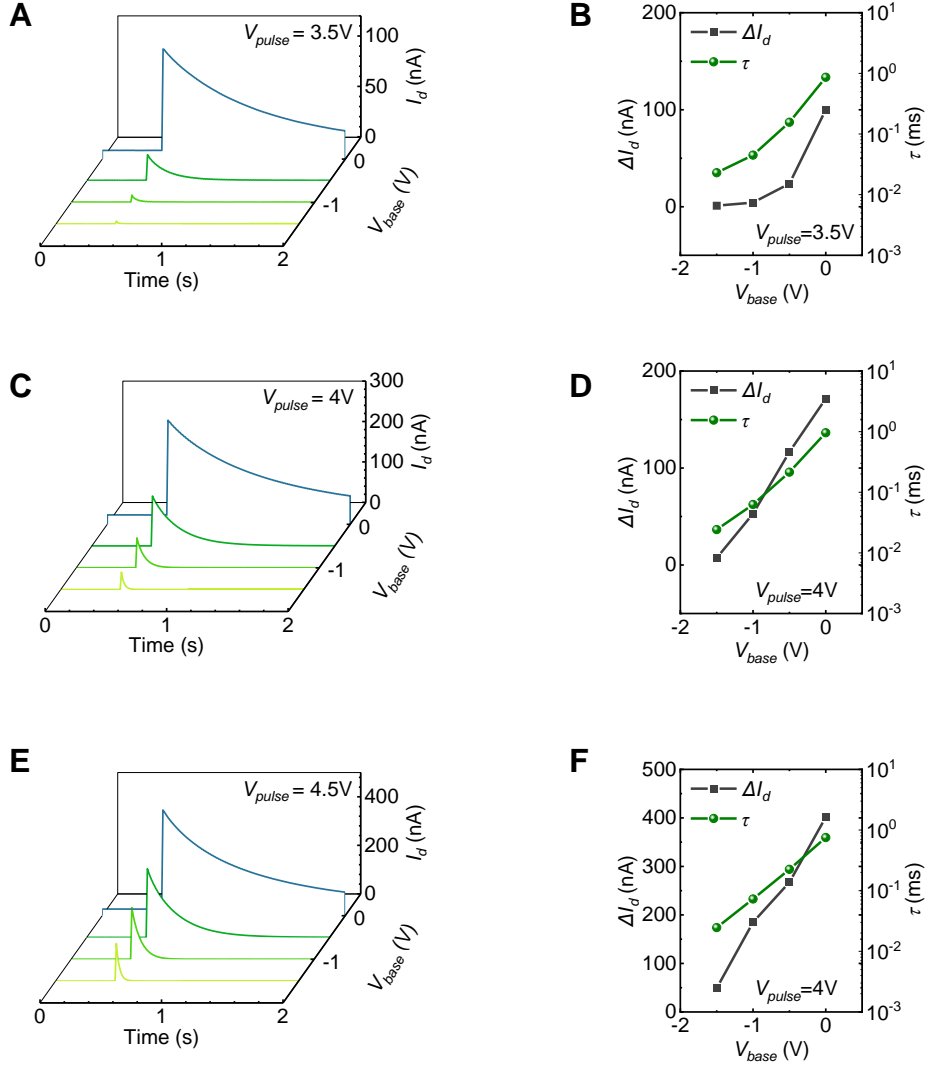
**Fig. S5 Bandgap diagram illustrating the redistribution process of electrons and relaxation process in ZnO.** The energy levels of defects in ZnO have been widely researched, which points out that V<sub>O</sub> in ZnO is negative U center (47-49). In other words, they are only stable in the form of V<sub>O</sub> or V<sub>O</sub><sup>2+</sup>, while thermodynamically unstable in the form of V<sub>O</sub><sup>+</sup>. As a result, in this analysis, we only consider the transition between V<sub>O</sub> and V<sub>O</sub><sup>2+</sup>. The above bandgap diagram schematically shows the 4 stages of charge redistribution. When positive  $V_g$  is applied, the electrons are activated from V<sub>O</sub> traps and swept towards the gate. After that, these electrons are re-trapped by V<sub>O</sub> with higher energy levels, since the lower energy levels are already occupied. In the relaxation process, the electrons will be gradually activated from a high energy level and transferred to a lower energy level, as a joint consequence of thermal movement and  $V_{base}$ . Finally, the initial distribution is recovered and so is the fading gate-modulating effect induced by V<sub>O</sub><sup>2+</sup>.



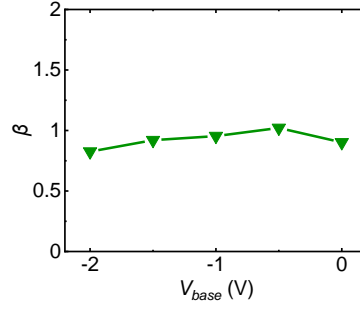
**Fig. S6 The accumulated gate charge during a DC double sweep.** In this double sweep, the gate voltage is swept from  $V_g = -4$  V to  $V_g = 4$  V and back to  $V_g = -4$  V. By integrating the gate current ( $I_g$ ) with time, a closed hysteresis of gate charge ( $Q_g$ ) is shown, reflecting the charge redistribution process and recovering process. The hysteresis shows that when  $V_g$  sweeps towards positive, positive charge will accumulate at the gate, and vice versa. This is in accordance with our aforementioned analysis that extra amounts of electrons will be trapped near the gate when positive  $V_g$  is applied. When a full cycle is performed, the  $Q_g$  returns to the original value, suggesting that no persisting alteration of the characteristics is caused by the sweep.



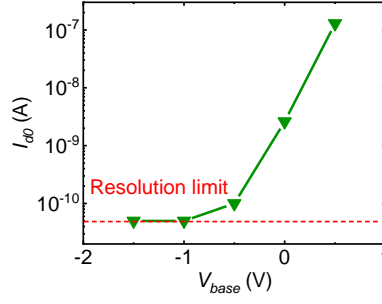
**Fig. S7 Dependence of  $\tau$  on  $V_{base}$ .** 50 repeated measurements were made for each  $V_{base}$  @ pulse width=20  $\mu$ s,  $V_{pulse}$ =3.5 V,  $V_d$ =0.1 V. As  $V_{base}$  changes from -3V to 1V,  $\tau$  rises from 91.7  $\mu$ s to 807 ms. The modulation range reaches nearly  $10^4$ . To investigate the cycle-to-cycle uniformity, we denote the standard deviation of  $\tau$  calculated from the 50 repeated measurements as  $\sigma(\tau)$ . Good uniformity is indicated by a low  $\sigma(\tau)$  level, which is always smaller than the value of  $\tau$  in 1 to 2 orders of magnitude. This result shows the modulation method to be effective and reliable.



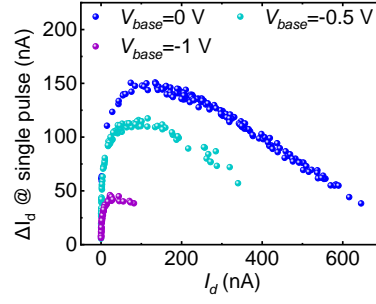
**Fig. S8 Modulation of the FM behavior by  $V_{base}$  in multiple  $V_{pulse}$  conditions.** The response of the same device to a single voltage pulse with  $V_{pulse}=3.5\text{ V}$ ,  $V_{pulse}=4\text{ V}$ , and  $V_{pulse}=4.5\text{ V}$  are measured respectively, showing the unified modulation effect. In all the above cases,  $\Delta I_d$  and  $\tau$  increase as  $V_{base}$  increases. The dependence relationship is quantitatively different but qualitatively identical. This indicates the modulation method with  $V_{base}$  is applicable for different input situations. All the measurements are conducted at pulse width= $20\text{ }\mu\text{s}$ ,  $V_d=0.1\text{ V}$ .



**Fig. S9 Dependence of  $\beta$  on  $V_{base}$ .** We illustrate the  $\beta$ - $V_{base}$  dependent relationship with the experiment data in Fig. 3G. As shown in the figure, no obvious variation in  $\beta$  can be observed among different  $V_{base}$  levels. This indicates that  $\beta$  is independent of our modulation method with  $V_{base}$ . Therefore, we can focus on  $\tau$  when analyzing FM characteristics. The measurement is conducted at pulse width=20  $\mu$ s,  $V_{pulse}$ =4 V,  $V_d$ =0.1 V.

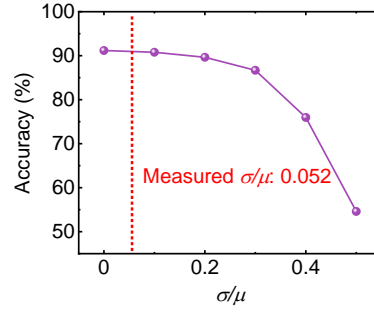


**Fig. S10 Dependence of initial current state ( $I_{d0}$ ) on  $V_{base}$ .**  $I_{d0}$ - $V_{base}$  relationship is measured at  $V_{base}=0.5$  V, 0 V, -0.5 V, -1 V and  $V_d=0.5$  V. The experiment indicates that the initial drain current  $I_{d0}$  decreases as  $V_{base}$  gets more negative (For  $V_{base}<-0.5$  V, the  $I_{d0}$  value stays unchanged only because it has exceeded the resolution limit of the measurement method). The lowered current is resulted from the suppression of carrier density by negative gate voltage.

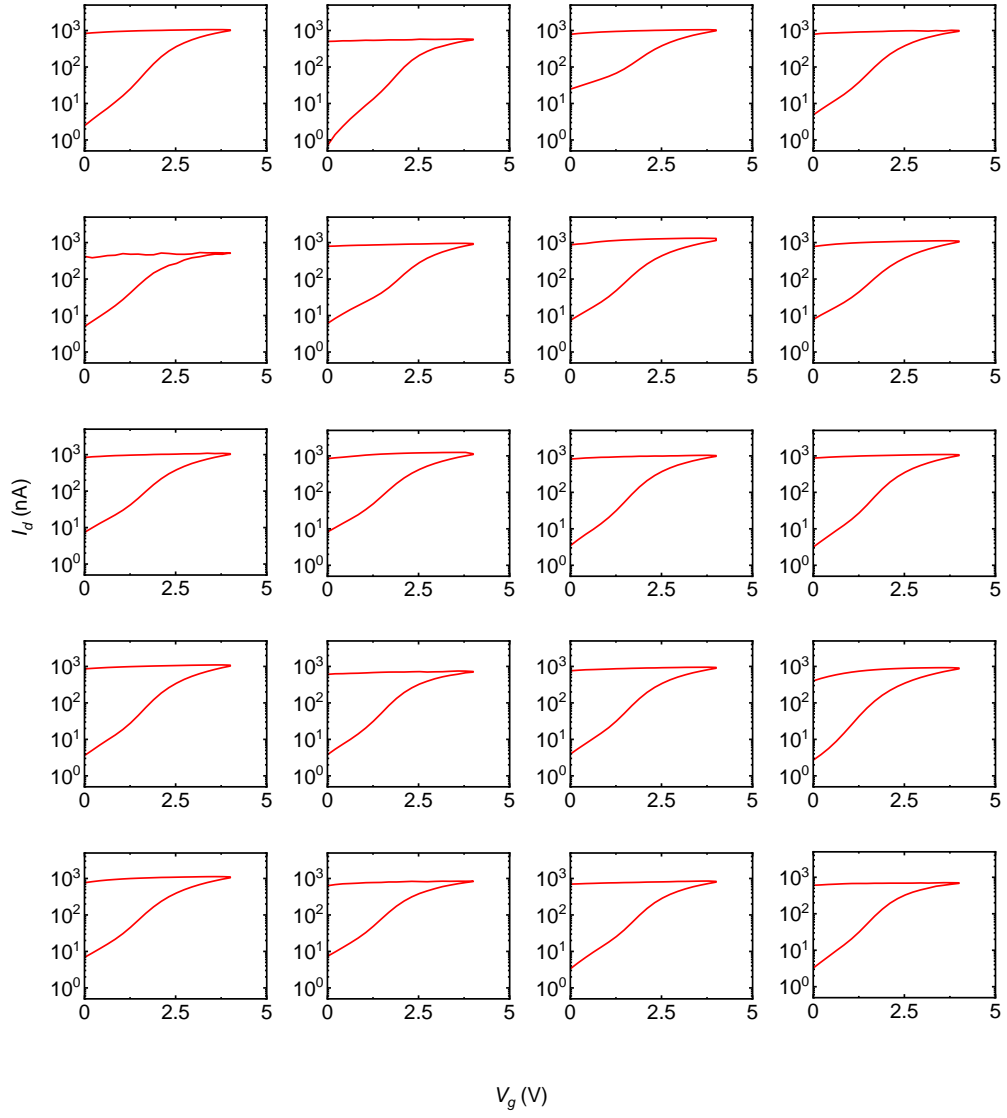


**Fig. S11 Dependence of activation process on  $V_{base}$ .** The activation process of various  $V_{base}$  is measured at pulse width=10  $\mu$ s,  $V_{pulse}$ =4 V,  $V_d$ =0.5 V. The increment of drain current  $\Delta I_d$  upon a single voltage pulse is used to represent the activation effect. It is observed that the  $\Delta I_d$  will decrease as  $V_{base}$  gets more negative, while the overall  $\Delta I_d$ - $I_d$  relationship is qualitatively similar for all diverse  $V_{base}$ . The differences in the initial state and activation process that are reflected as the magnitude of the readout current have only slight influence on the following classification, since the readout current will perform the  $V_{base}$ -specific normalization operation before being fed into the FCN layers. (see Materials and Methods).

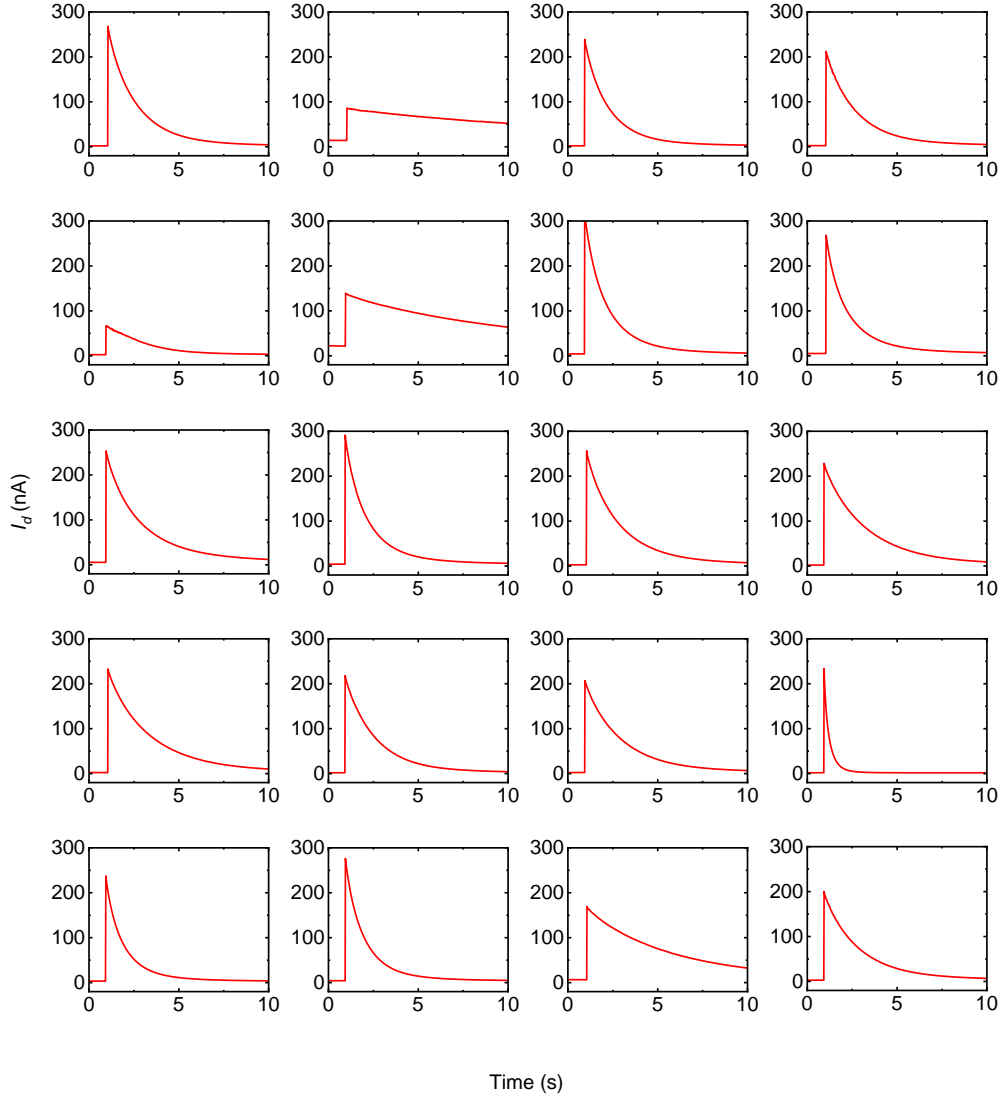




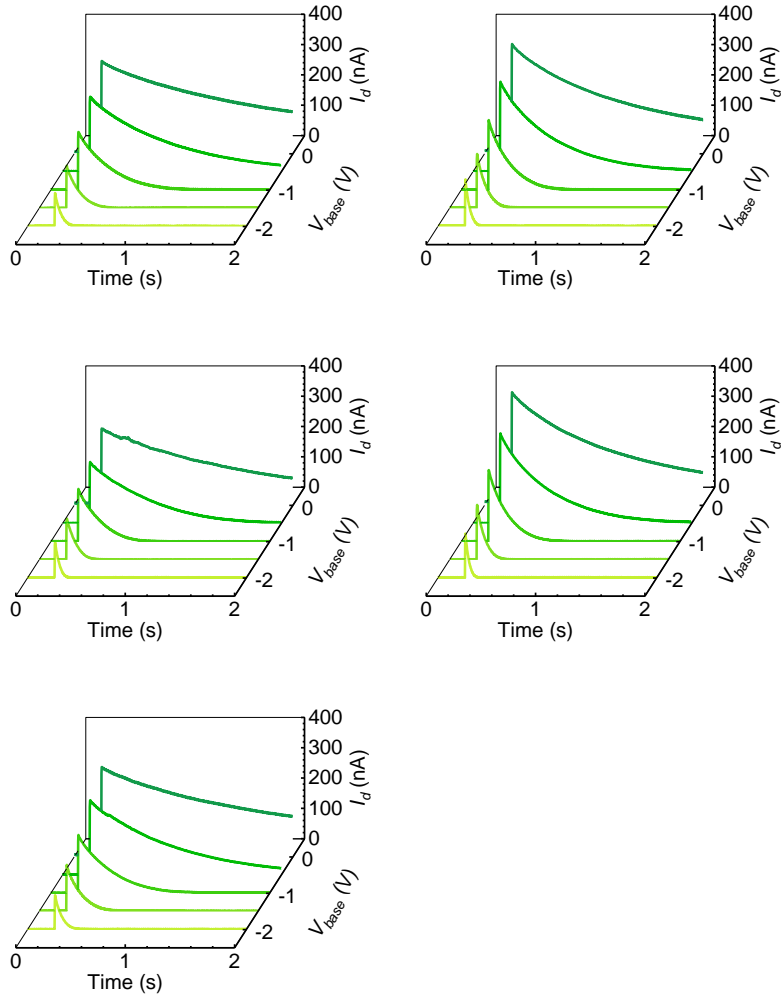
**Fig. S12 Evaluation of the impact of cycle-to-cycle variation on recognition accuracy.** We have simulated the accuracy of the HAR task in different noise levels of the readout drain current. The accuracy firstly decreases slightly with the increasing of cycle-to-cycle variation, and falls drastically when the  $\sigma/\mu$  is above 0.2. For our devices, the noise resulting from cycle-to-cycle variation is  $\sigma/\mu = 0.052$ , which will not lead to a distinct accuracy degradation.



**Fig. S13 Device-to-device variation of the  $V_g$ -sweep behavior.** We randomly chose an area on the wafer and tested several continuously distributed devices. The  $V_g$ -sweep behavior of 20 well-functioning devices has shown qualitatively identical features. The yield is estimated to be 71.43% since the total number of tested devices is 28. The measurement parameter is  $V_d=0.1$  V.

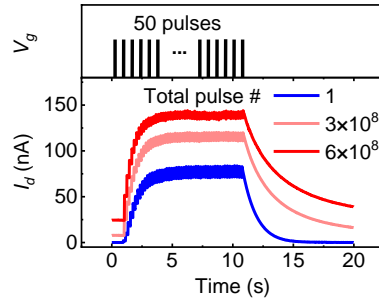


**Fig. S14 Device-to-device variation of the single-pulse response.** The single pulse measurement of the 20 selected devices in Fig. S13. Similar activation and relaxation behaviors are observed for all devices. The measurement parameters are pulse width=10  $\mu$ s,  $V_{base}$ =-1 V,  $V_{pulse}$ =4 V,  $V_d$ =0.1 V.



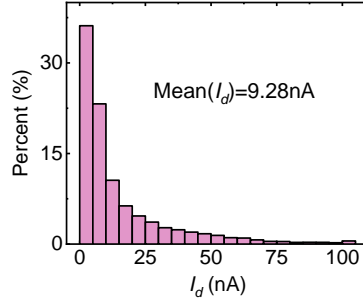
**Fig. S15 Device-to-device variation of the modulation of fading memory behavior.**

We measured 5 randomly chosen devices under different  $V_{base}$  levels. The modulation scheme is confirmed to apply to all devices. The measurement parameters are pulse width=20  $\mu$ s,  $V_{pulse}$ =4 V,  $V_d$ =0.1 V.

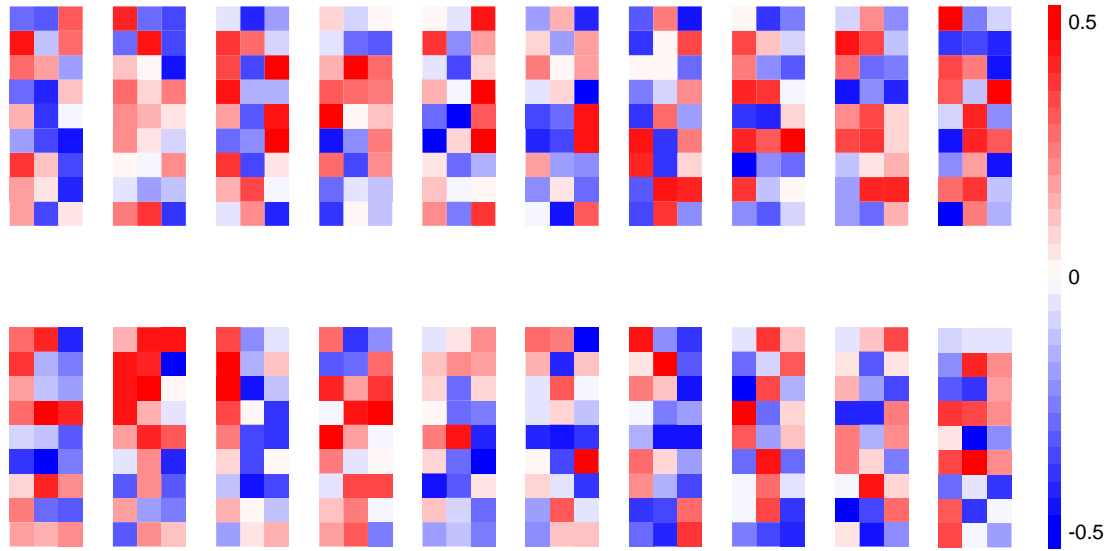


**Fig. S16 Degradation of the device characteristics in the endurance measurement.**

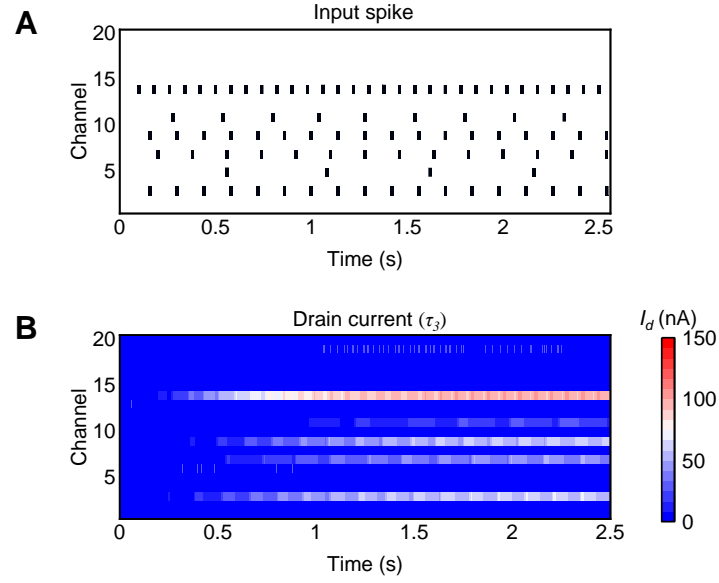
The responsive curve of 50 pulses begins to shift away as the total number of applied pulses continues to accumulate. Specifically, both the initial value of current and  $\tau$  have gradually increased, yielding a higher responsive current. Based on precedent research of positive-bias temperature instability (PBTI) effect (57, 58), we speculate the degradation mechanism to be the creation of electron traps in the oxide layer. When high positive  $V_g$  is applied, electrons in ZnO are injected into the oxide layer through Fowler-Nordheim tunneling, leading to the hot-carrier-induced creation of defects in  $\text{Al}_2\text{O}_3$ . These newly created defects function as deep level traps. The electrons trapped in those traps will have less de-trapping probability because of higher energy barrier, which results in a longer  $\tau$  in the relaxation process. Moreover, as the pulse number keeps increasing, part of the electrons will reside and accumulate in the  $\text{Al}_2\text{O}_3$  defects rather than being swept back, which will result in the increased initial current. Based on the analysis, we propose that one possible solution for improving the endurance characteristics is to optimize the crystallization quality of high-k dielectric layer, so as to reduce the amount of newly created defects (59).



**Fig. S17 Energy efficiency evaluation of the reservoir computing system.** We use the data collected from the HAR experiment to analyze the energy efficiency of the reservoir system. The analysis is based on the recorded reservoir states, i.e., the measured drain current of FM TFT during the hardware inference. The inference was conducted with samples from the test dataset, where all categories of input signals were included. We plot the histogram of all the sampled drain current throughout inference. It can be seen that the majority of the measured current value is less than 20 nA, with the mean value being 9.28 nA. This is a very low operation current compared to counterpart works. Since the gate current is sufficiently low as shown in Fig. S6, the drain current and source current constitute the majority of power consumption. Therefore, we concluded that this hardware RC system has quite high energy efficiency.

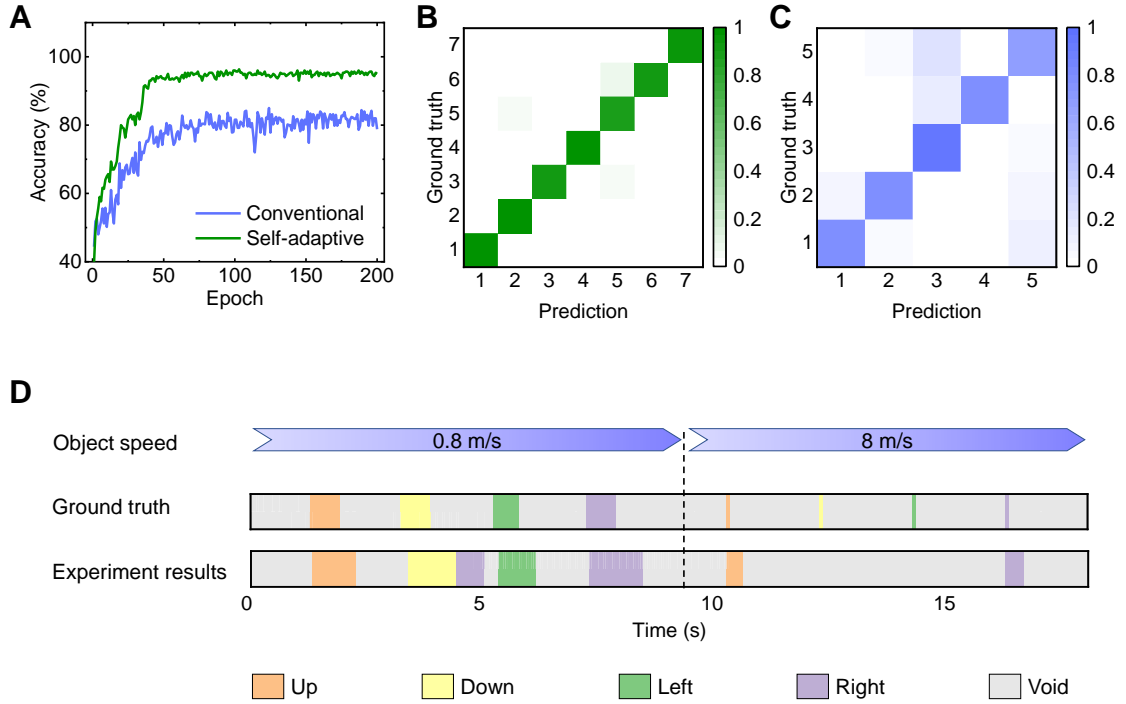


**Fig. S18 The 20 convolution kernels applied in the HAR task.** The 20 convolution kernels are of the size of  $9 \times 3$ , within which each element is randomly generated from the range of -0.5 to 0.5.



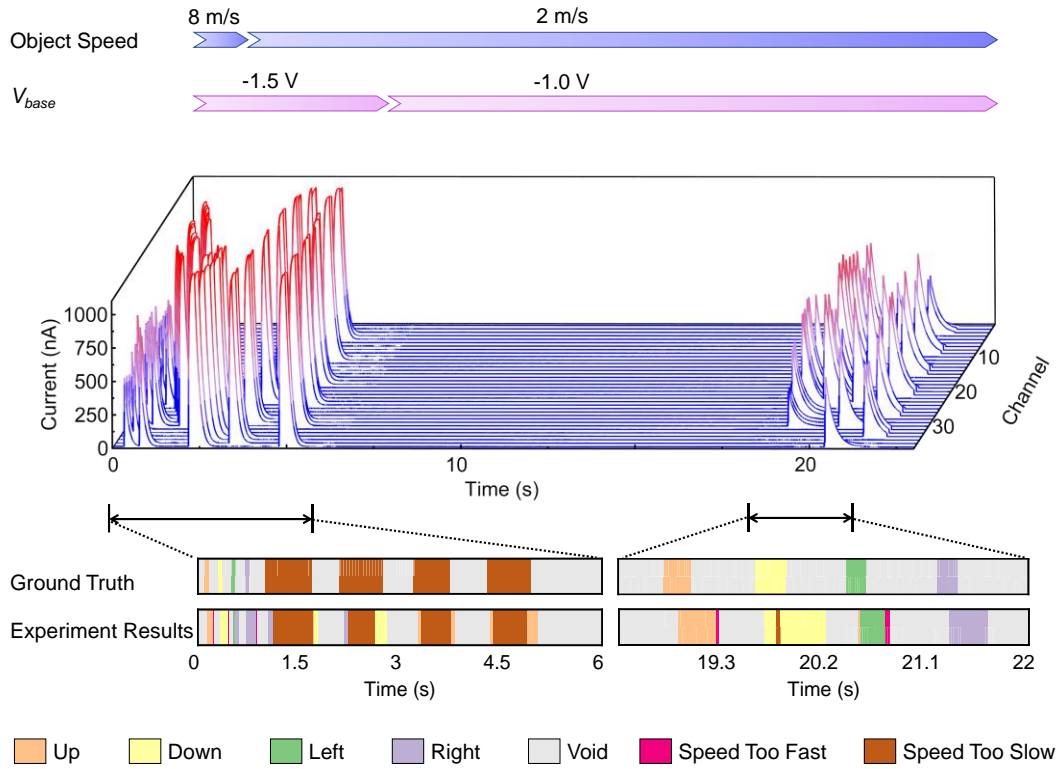
**Fig. S19 Intermediate results of the HAR task.** (A) The spike sequence corresponding to a selected sample labeled “laying”. The spike sequence is generated by convolution and binarization. It will then be converted to the gate voltage as input of the reservoir. (B) shows the experimentally measured response of drain current to the input in (A). It is observed that the value of drain current will rise on the stimulus of voltage spike, and then gradually decay. 30 points will be sampled for each channel, submitting 600 nodes in total for the subsequent FCN.



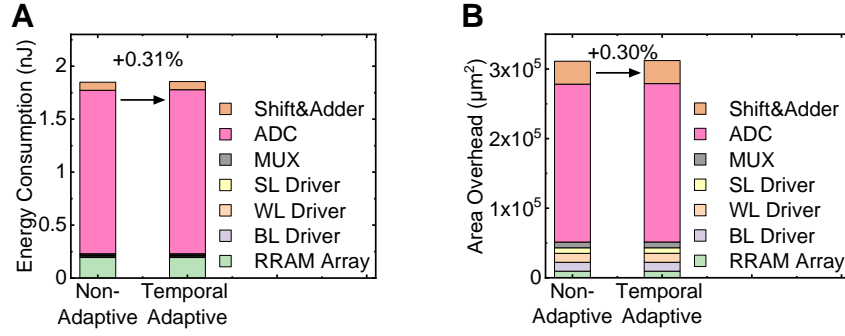


**Fig. S20 Comparison of conventional RC and self-adaptive RC with closed-loop in the real-time recognition task.** In order to prove the advantage of self-adaptive reservoir with closed-loop structure as we introduced in Fig. 6, we applied a conventional temporal invariant reservoir system to the same real-time recognition task for comparison. The conventional reservoir system classifies the input series into 5 categories, namely moving up, down, left, right, and void. **(A)** shows the comparison between conventional and self-adaptive reservoirs. As the training continues, the accuracy of both reservoirs first rises and then converges within 100 epochs. The self-adaptive reservoir system outperforms the conventional reservoir not only in convergence speed but also in the ultimate accuracy, which is 94.67% @ 7 categories versus 78.8% @ 5 categories. The prominent difference is caused by the strategy in tackling multi-timescale information. In this simulation, we constructed the training and testing dataset by including equal number of signals of objects moving at 3 speeds (0.08 m/s, 0.8 m/s, 8 m/s). The characteristics of the reservoir are extracted from data measured at  $V_{base} = -0.5$  V, which is most in favor of the middle speed (0.8 m/s) and can best balance the slow and fast cases. The conventional RC is designed to classify the input samples regardless of the timescale. Its poor accuracy indicates the fact that RC cannot function well when encountering incompatible timescales. In contrast, the self-

adaptive RC is designed to only classify the input with matched timescale, and the mismatching event itself. Its superior accuracy demonstrates that RC is intrinsically better at identifying temporal mismatch, rather than conducting classification in those conditions. **(B)** and **(C)** shows the confusion matrix of the above two reservoir systems. **(D)** shows the experimental result of applying the conventional reservoir system to the varying-speed condition. During the experiment the  $V_{base}$  is kept at -0.5 V. When the speed is 0.8 m/s, the RC system can output a relatively accurate outcome. While if the speed increases to 8 m/s, the conventional RC system begins to show obvious mistakes in classifying moving directions, which cannot compete with the self-adaptive reservoir system in Fig. 6G. The difference is caused by the same reason as discussed in **(A)**. The comparison validates that the self-adaptive RC is an effective solution in dealing with temporal compound input signals.



**Fig. S21 Experimental demonstration of closed-loop reservoir in a slow-speed adaptation case.** The speed of the moving object was transferred from 8 m/s to 2 m/s. The closed-loop reservoir has responded to this transition by converting  $V_{base}$  from -1.5 V to -1 V. The closed-loop RC system can accurately recognize the moving direction and the speed mismatching, which has confirmed its capability in the speed-decreasing situation.



**Fig. S22 Evaluation of the energy and area overhead of the FCN part.** The evaluated expense of the temporal adaptability almost can be ignored based on the simulation of 2000 samples in the motion detection task at the recognition rate of 200 Hz, as shown in Fig. S22. Since the temporal adaptive RC shares the same reservoir structure as the non-adaptive RC, the reservoir part has no additional overhead. Thus, we do not consider the energy consumption and area overhead of the reservoir part in the evaluation. We evaluate the overhead of the subsequent hardware FCN by considering a memristor-based neural processing unit. Several of the most dominant parts are included: the memristor array along with the driver of the bit line, word line, and source line; column-parallel ADC and the column-selection MUX; and a digital adder that yields the final output. The overall energy consumption and area of the hardware FCN are calculated according to the metrics listed in Table S1, extracted from a recently reported memristor-based processing unit (60). In the FCN part, the self-adaptive RC system requires structure alteration for the last FCN layer, resulting in the extra overhead of the memristor array, source line driver, MUX, ADC, and digital adder. The exact expense is calculated to be +0.31% and +0.30% for energy consumption and area overhead, respectively, as shown in (A) and (B). The results indicate that a little more energy consumption and area overhead can bring the temporal adaptive capability.

**Table S1 Metrics of energy and area overhead of each circuitry module used for evaluation.** We evaluate the energy overhead of the subsequent hardware FCN by considering a memristor-based neural processing unit. Several most dominant parts are included: the memristor array along with the driver of the bit line, word line, and source line; column-parallel ADC and the column-selection MUX; and digital adder that yields the final output. The overall energy & area overhead of the hardware FCN is calculated according to metrics of a recently reported memristor-based processing unit (60).

Module	Energy	Area
Array	1.38 fJ/(col*row)	0.07 $\mu\text{m}^2$ /(col*row)
BL driver	15.55 fJ/row	16.44 $\mu\text{m}^2$ /row
WL driver	15.55 fJ/row	16.44 $\mu\text{m}^2$ /row
SL driver	0.86 fJ/col	13.17 $\mu\text{m}^2$ /col
MUX	18.51 fJ/col	13.61 $\mu\text{m}^2$ /col
ADC	2550.00 fJ/col	375.00 $\mu\text{m}^2$ /col
Shifter & adder	127.03 fJ/col	54.51 $\mu\text{m}^2$ /col

## REFERENCES AND NOTES

1. Y. Lecun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
2. O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. E. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: A survey. *Heliyon* **4**, e00938 (2018).
3. J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, A. Troncoso, Deep learning for time series forecasting: A survey. *Big Data* **9**, 3–21 (2021).
4. J. Cui, F. An, J. Qian, Y. Wu, L. L. Sloan, S. Pidaparthi, J. M. Zuo, Q. Cao, CMOS-compatible electrochemical synaptic transistor arrays for deep learning accelerators. *Nat. Electron.* **6**, 292–300 (2023).
5. J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554–2558 (1982).
6. P. J. Werbos, Backpropagation through time: What it does and how to do it. *Proc. IEEE* **78**, 1550–1560 (1990).
7. R. Pascanu, D. Tour, T. Mikolov, D. Tour, On the difficulty of training recurrent neural networks, in *Proceedings of the 30th International Conference on Machine Learning* (PMLR, 2013), pp. 1310–1318.
8. Y. Bengio, N. Boulanger-Lewandowski, R. Pascanu, Advances in optimizing recurrent networks, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 2013), pp. 8624–8628.
9. W. Maass, T. Natschlager, H. Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
10. K. Holland, Report from Nurse Education Tomorrow 2001: 12th Annual International Participative Conference—For Education in Health Care, Grey College, University of Durham, UK. 7–9 September 2001. *Nurse Educ. Pract.* **1**, 221–223 (2001).

11. M. Lukoševičius, H. Jaeger, Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127–149 (2009).
12. Z. Lu, B. R. Hunt, E. Ott, Attractor reconstruction by machine learning. *Chaos* **28**, 061104 (2018).
13. J. A. Platt, A. Wong, R. Clark, S. G. Penny, H. D. I. Abarbanel, Robust forecasting using predictive generalized synchronization in reservoir computing. *Chaos* **31**, 123118 (2021).
14. L. Appeltant, M. C. Soriano, G. Van Der Sande, J. Danckaert, S. Massar, J. Dambre, G. Van Der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, I. Fischer, Information processing using a single dynamical node as complex system. *Nat. Commun.* **2**, 468 (2011).
15. B. Schrauwen, D. Verstraeten, J. Van Campenhout, An overview of reservoir computing: Theory, applications and implementations, in *Proceedings of the 15th European Symposium on Artificial Neural Networks* (ESANN, 2007), pp. 471–482.
16. D. J. Gauthier, E. Bollt, A. Griffith, W. A. S. Barbosa, Next generation reservoir computing. *Nat. Commun.* **12**, 5564 (2021).
17. M. Cucchi, S. Abreu, G. Ciccone, D. Brunner, H. Kleemann, Hands-on reservoir computing: A tutorial for practical implementation. *Neuromorph. Comput. Eng.* **2**, 032002 (2022).
18. H. Hauser, A. J. Ijspeert, R. M. Fuchslin, R. Pfeifer, W. Maass, Towards a theoretical foundation for morphological computation with compliant bodies. *Biol. Cybern.* **105**, 355–370 (2011).
19. G. Urbain, J. Degraeve, B. Carette, J. Dambre, F. Wyffels, Morphological properties of mass-spring networks for optimal locomotion learning. *Front. Neurobot.* **11**, 16 (2017).
20. P. Antonik, N. Marsal, D. Brunner, D. Rontani, Human action recognition with a large-scale brain-inspired photonic computer. *Nat. Mach. Intell.* **1**, 530–537 (2019).

21. P. Antonik, N. Marsal, D. Brunner, D. Rontani, Bayesian optimisation of large-scale photonic reservoir computers. *Cognit. Comput.* **15**, 1452–1460 (2023).
22. G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, A. Hirose, Recent advances in physical reservoir computing: A review. *Neural Netw.* **115**, 100–123 (2019).
23. C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, W. D. Lu, Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **8**, 2204 (2017).
24. Z. Qi, L. Mi, H. Qian, W. Zheng, Y. Guo, Y. Chai, Physical reservoir computing based on nanoscale materials and devices. *Adv. Funct. Mater.* **33**, 2306149 (2023).
25. R. Midya, Z. Wang, S. Asapu, X. Zhang, M. Rao, W. Song, Y. Zhuo, N. Upadhyay, Q. Xia, J. J. Yang, Reservoir computing using diffusive memristors. *Adv. Intell. Syst.* **1**, 1900084 (2019).
26. Y. Zhong, J. Tang, X. Li, B. Gao, H. Qian, H. Wu, Dynamic memristor-based reservoir computing for high-efficiency temporal signal processing. *Nat. Commun.* **12**, 408 (2021).
27. W. Sun, W. Zhang, J. Yu, Y. Li, Z. Guo, J. Lai, D. Dong, X. Zheng, F. Wang, S. Fan, X. Xu, D. Shang, M. Liu, 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)* (IEEE, 2022), pp. 222–223.
28. L. Sun, Z. Wang, J. Jiang, Y. Kim, B. Joo, S. Zheng, S. Lee, W. J. Yu, B. S. Kong, H. Yang, In-sensor reservoir computing for language learning via two-dimensional memristors. *Sci. Adv.* **7**, eabg1455 (2021).
29. H. Li, S. Wang, X. Zhang, W. Wang, R. Yang, Z. Sun, W. Feng, P. Lin, Z. Wang, L. Sun, Y. Yao, Memristive crossbar arrays for storage and computing applications. *Adv. Intell. Syst.* **3**, 2100017 (2021).



30. J. Yu, Y. Li, W. Sun, W. Zhang, Z. Gao, D. Dong, Z. Yu, Y. Zhao, J. Lai, Q. Ding, Q. Luo, C. Dou, Q. Zuo, Y. Zhao, S. Chen, R. Zou, H. Chen, Q. Wang, H. Lv, X. Xu, D. Shang, M. Liu, Energy efficient and robust reservoir computing system using ultrathin (3.5 nm) ferroelectric tunneling junctions for temporal data learning, in *2021 Symposium on VLSI Technology* (IEEE, 2021), pp. 2021–2022.
31. E. Nako, K. Toprasertpong, R. Nakane, M. Takenaka, S. Takagi, Experimental demonstration of novel scheme of HZO/Si FeFET reservoir computing with parallel data processing for speech recognition, in *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)* (IEEE, 2022), pp. 220–221.
32. M. Tang, J. Mei, X. Zhan, C. Wang, J. Chai, H. Xu, X. Wang, J. Wu, J. Chen, Fully ferroelectric-fets reservoir computing network for temporal and random signal processing. *IEEE Trans. Electron Devices* **70**, 3372–3377 (2023).
33. G. Milano, G. Pedretti, K. Montano, S. Ricci, S. Hashemkhani, L. Boarino, D. Ielmini, C. Ricciardi, *In materia* reservoir computing with a fully memristive architecture based on self-organizing nanowire networks. *Nat. Mater.* **21**, 195–202 (2022).
34. R. K. Daniels, J. B. Mallinson, Z. E. Heywood, P. J. Bones, M. D. Arnold, S. A. Brown, Reservoir computing with 3D nanowire networks. *Neural Netw.* **154**, 122–130 (2022).
35. R. Fang, W. Zhang, K. Ren, P. Zhang, X. Xu, Z. Wang, D. Shang, In-materio reservoir computing based on nanowire networks: Fundamental, progress, and perspective. *Mater. Futures* **2**, 022701 (2023).
36. J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, J. Grollier, Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
37. M. Riou, F. Abreu Araujo, J. Torrejon, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, J. Grollier, Neuromorphic

- computing through time-multiplexing with a spin-torque nano-oscillator, in *2017 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2018), pp. 36.3.1–36.3.4.
38. R. Nakane, G. Tanaka, A. Hirose, Reservoir computing with spin waves excited in a garnet film. *IEEE Access* **6**, 4462–4469 (2018).
39. D. Joksas, A. AlMutairi, O. Lee, M. Cubukcu, A. Lombardo, H. Kurebayashi, A. J. Kenyon, A. Mehonic, Memristive, spintronic, and 2D-materials-based devices to improve and complement computing hardware. *Adv. Intell. Syst.* **4**, 2200068 (2022).
40. X. Zhu, Q. Wang, W. D. Lu, Memristor networks for real-time neural activity analysis. *Nat. Commun.* **11**, 2439 (2020).
41. Y. H. Jang, W. Kim, J. Kim, K. S. Woo, H. J. Lee, J. W. Jeon, S. K. Shim, J. Han, C. S. Hwang, Time-varying data processing with nonvolatile memristor-based temporal kernel. *Nat. Commun.* **12**, 5727 (2021).
42. J. Moon, W. Ma, J. H. Shin, F. Cai, C. Du, S. H. Lee, W. D. Lu, Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat. Electron.* **2**, 480–487 (2019).
43. Y. Zhong, J. Tang, X. Li, X. Liang, Z. Liu, Y. Li, Y. Xi, P. Yao, Z. Hao, B. Gao, H. Qian, H. Wu, A memristor-based analogue reservoir computing system for real-time and power-efficient signal processing. *Nat. Electron.* **5**, 672–681 (2022).
44. K. Liu, B. Dang, T. Zhang, Z. Yang, L. Bao, L. Xu, C. Cheng, R. Huang, Y. Yang, Multilayer reservoir computing based on ferroelectric  $\alpha$ - $\text{In}_2\text{Se}_3$  for hierarchical information processing. *Adv. Mater.* **34**, e2108826 (2022).
45. K. Liu, T. Zhang, B. Dang, L. Bao, L. Xu, C. Cheng, Z. Yang, R. Huang, Y. Yang, An optoelectronic synapse based on  $\alpha$ - $\text{In}_2\text{Se}_3$  with controllable temporal dynamics for multimode and multiscale reservoir computing. *Nat. Electron.* **5**, 761–773 (2022).

46. N. Jiang, J. Tang, W. Zhang, Y. Li, N. Li, X. Li, X. Chen, R. Fang, Z. Guo, F. Wang, J. Wang, Z. Li, C. He, G. Zhang, Z. Wang, D. Shang, Bioinspired in-sensor reservoir computing for self-adaptive visual recognition with two-dimensional dual-mode phototransistors. *Adv. Opt. Mater.* **11**, 2300271 (2023).
47. D.M. Hofmann, D. Pfisterer, J. Sann, B.K. Meyer, R. Tena-Zaera, V. Munoz-Sanjose, T. Frank, G. Pensl, Properties of the oxygen vacancy in ZnO. *Appl. Phys A* **88**, 147–151 (2007).
48. J. Wang, R. Chen, L. Xiang, S. Komarneni, Synthesis , properties and applications of ZnO nanomaterials with oxygen vacancies : A review. *Ceram. Int.* **44**, 7357–7377 (2018).
49. V. Gurylev, T. P. Perng, Defect engineering of ZnO: Review on oxygen and zinc vacancies. *J. Eur. Ceram. Soc.* **41**, 4977–4996 (2021).
50. J. C. Phillips, Stretched exponential relaxation in molecular and electronic glasses. *Rep. Prog. Phys.* **59**, 1133–1207 (1996).
51. D. Anguita, A. Ghio, L. Oneto, X. Parra, J. L. Reyes-Ortiz, Energy efficient smartphone-based activity recognition using fixed-point arithmetic. *J. Univers. Comput. Sci.* **19**, 1295–1314 (2013).
52. S. Wang, H. Chen, W. Zhang, Y. Li, D. Wang, S. Shi, Y. Zhao, K. C. Loong, X. Chen, Y. Dong, Y. Zhang, Y. Jiang, C. Furqan, J. Chen, Q. Wang, X. Xu, G. Wang, H. Yu, D. Shang, Z. Wang, Convolutional echo-state network with random memristors for spatiotemporal signal classification. *Adv. Intell. Syst.* **4**, 2200027 (2022).
53. J. Moon, Y. Wu, W. D. Lu, Hierarchical architectures in reservoir computing systems. *Neuromorph. Comput. Eng.* **1**, 014006 (2021).
54. J. Chen, Z. Zhou, B. J. Kim, Y. Zhou, Z. Wang, T. Wan, J. Yan, J. Kang, J.-H. Ahn, Y. Chai, Optoelectronic graded neurons for bioinspired in-sensor motion perception. *Nat. Nanotechnol.* **18**, 882–888 (2023).

55. A. Daus, C. Vogt, N. Münzenrieder, L. Petti, S. Knobelspies, G. Cantarella, M. Luisier, G. A. Salvatore, G. Tröster, Positive charge trapping phenomenon in n-channel thin-film transistors with amorphous alumina gate insulators. *J. Appl. Phys.* **120**, 244501 (2016).
56. S. J. Ding, X. Wu, Superior atomic layer deposition technology for amorphous oxide semiconductor thin-film transistor memory devices. *Chem. Mater.* **32**, 1343–1357 (2020).
57. A. Chasin, J. Franco, K. Triantopoulos, H. Dekkers, N. Rassoul, A. Belmonte, Q. Smets, S. Subhechha, D. Claes, M. J. Van Setten, J. Mitard, R. Delhougne, V. Afanas'Ev, B. Kaczer, G. S. Kar, Understanding and modelling the PBTI reliability of thin-film IGZO transistors, in *2021 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2021) pp. 31.1.1–31.1.4.
58. J. Rhee, S. Choi, H. Kang, J. Y. Kim, D. Ko, G. Ahn, H. Jung, S. J. Choi, D. Myong Kim, D. H. Kim, The electron trap parameter extraction-based investigation of the relationship between charge trapping and activation energy in IGZO TFTs under positive bias temperature stress. *Solid State Electron.* **140**, 90–95 (2018).
59. H. J. Lim, Y. Kim, I. Sang Jeon, J. Yeo, B. Im, S. Hong, B. H. Kim, S. W. Nam, H.-K. Kang, E. S. Jung, Impact of the crystallization of the high-k dielectric gate oxide on the positive bias temperature instability of the n-channel metal-oxide-semiconductor field emission transistor. *Appl. Phys. Lett.* **102**, 232909 (2013).
60. P. Yao, H. Wu, B. Gao, J. Tang, Q. Zhang, W. Zhang, J. J. Yang, H. Qian, Fully hardware-implemented memristor convolutional neural network. *Nature* **577**, 641–646 (2020).