

# In situ training of an in-sensor artificial neural network based on ferroelectric photosensors

Received: 23 July 2024

Accepted: 11 December 2024

Published online: 07 January 2025



Haipeng Lin<sup>1,5</sup>, Jiali Ou<sup>1,5</sup>, Zhen Fan<sup>1</sup>✉, Xiaobing Yan<sup>2</sup>✉, Wenjie Hu<sup>1</sup>, Boyuan Cui<sup>1</sup>, Jikang Xu<sup>2</sup>, Wenjie Li<sup>1</sup>, Zhiwei Chen<sup>1</sup>, Biao Yang<sup>2</sup>, Kun Liu<sup>1</sup>, Linyuan Mo<sup>1</sup>, Meixia Li<sup>1</sup>, Xubing Lu<sup>1</sup>, Guofu Zhou<sup>3</sup>, Xingsen Gao<sup>1</sup> & Jun-Ming Liu<sup>1,4</sup>

In-sensor computing has emerged as an ultrafast and low-power technique for next-generation machine vision. However, in situ training of in-sensor computing systems remains challenging due to the demands for both high-performance devices and efficient programming schemes. Here, we experimentally demonstrate the in situ training of an in-sensor artificial neural network (ANN) based on ferroelectric photosensors (FE-PSs). Our FE-PS exhibits self-powered, fast ( $<30\ \mu\text{s}$ ), and multilevel ( $>4$  bits) photoresponses, as well as long retention (50 days), high endurance ( $10^9$ ), high write speed (100 ns), and small cycle-to-cycle and device-to-device variations ( $\sim 0.66\%$  and  $\sim 2.72\%$ , respectively), all of which are desirable for the in situ training. Additionally, a bi-directional closed-loop programming scheme is developed, achieving a precise and efficient weight update for the FE-PS. Using this programming scheme, an in-sensor ANN based on the FE-PSs is trained in situ to recognize traffic signs for commanding a prototype autonomous vehicle. Moreover, this in-sensor ANN operates 50 times faster than a von Neumann machine vision system. This study paves the way for the development of in-sensor computing systems with in situ training capability, which may find applications in new data-streaming machine vision tasks.

Conventional von Neumann machine vision systems, composed of physically separated image sensing, memory, and processing units, are becoming time and energy inefficient in time-critical applications like autonomous driving. A promising solution to this problem is the recently emerging in-sensor computing technique<sup>1,2</sup>. In this technique, visual information is directly processed within image sensors, resulting in significantly reduced latency and energy consumption. So far a variety of in-sensor computing systems, mimicking both the low-level

(e.g., contrast enhancement and noise suppression)<sup>3–8</sup> and high-level (e.g., recognition and classification)<sup>9–19</sup> image processing functions of human visual systems, have been demonstrated. Among them, in-sensor artificial neural networks (ANNs), consisting of interconnected programmable photosensors with tunable photoresponsivities (weights), are of particular interest because they can realize real-time image sensing and recognition<sup>9–11,18,19</sup>. Such capability is highly demanded for the time-critical applications.

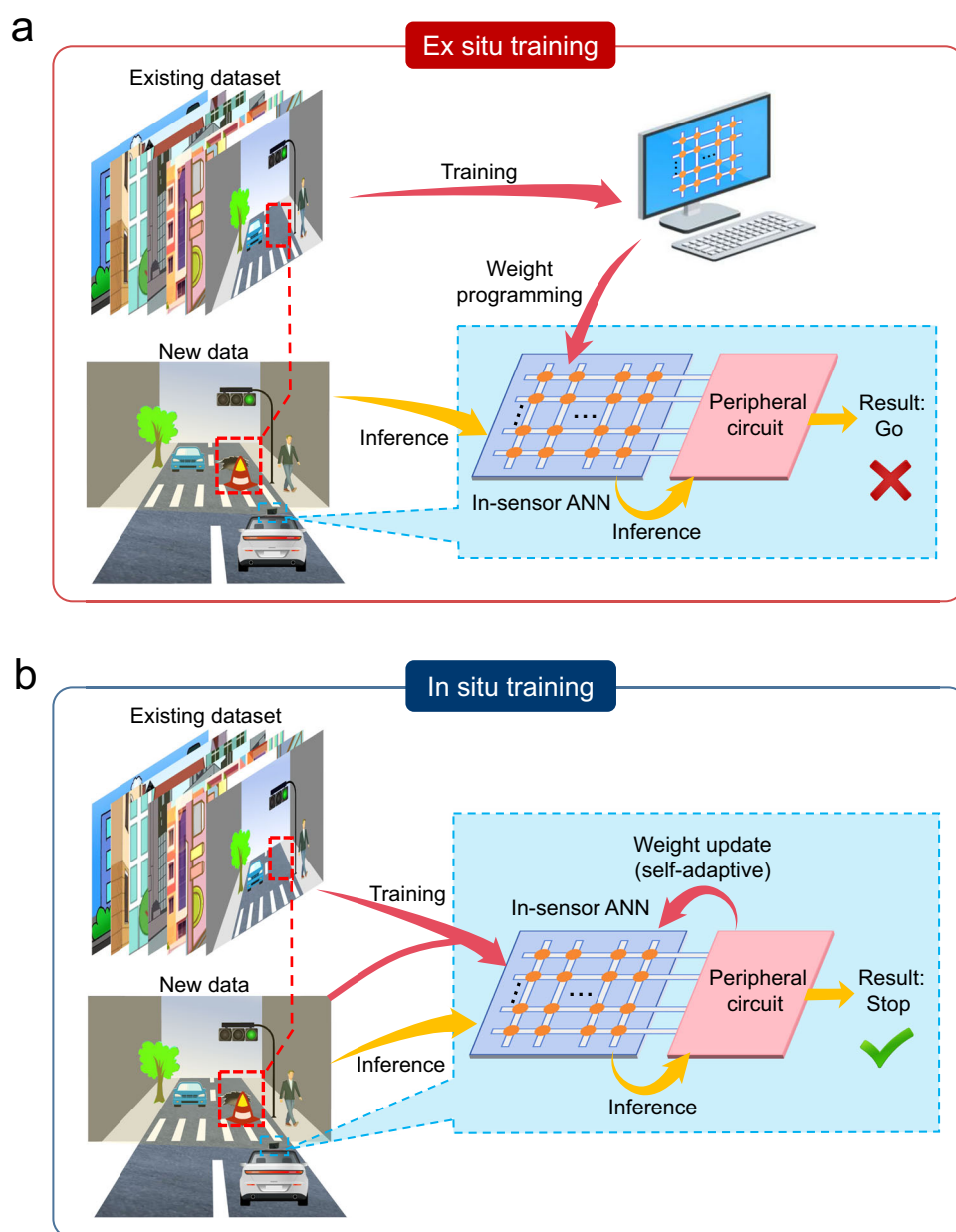
<sup>1</sup>Institute for Advanced Materials and Guangdong Provincial Key Laboratory of Optical Information Materials and Technology, South China Academy of Advanced Optoelectronics, South China Normal University, Guangzhou, China. <sup>2</sup>Key Laboratory of Brain-Like Neuromorphic Devices and Systems of Hebei Province, Hebei Key Laboratory of Photo-Electricity Information and Materials, Hebei University, Baoding, China. <sup>3</sup>National Center for International Research on Green Optoelectronics, South China Normal University, Guangzhou, China. <sup>4</sup>Laboratory of Solid State Microstructures and Innovation Center of Advanced Microstructures, Nanjing University, Nanjing, China. <sup>5</sup>These authors contributed equally: Haipeng Lin, Jiali Ou. ✉e-mail: [fanzhen@m.scnu.edu.cn](mailto:fanzhen@m.scnu.edu.cn); [yanxiaobing@ime.ac.cn](mailto:yanxiaobing@ime.ac.cn)

Notably, most previous studies implemented the in-sensor ANNs by using *ex situ* training<sup>17–22</sup>, where weights are computed by software models deployed on digital computers and, then, programmed into the hardware (Fig. 1a). By contrast, *in situ* training, where weights are updated directly in the hardware (Fig. 1b), was rarely implemented for in-sensor ANNs<sup>9,11</sup>. In fact, the *in situ* training is increasingly seen as a more efficient approach due to its advantages as follows<sup>23–27</sup>. First, the *in situ* training enhances the area and energy efficiencies of the hardware because it avoids the use of additional digital computers. In addition, the *in situ* training can tolerate some device non-idealities (e.g., stuck fault and device-to-device variation) owing to the self-adaptive weight adjustment, thus resulting in improved training performance. Besides, the *in situ* training empowers the system with an online learning capability, which is important for handling new data-streaming scenarios like autonomous driving.

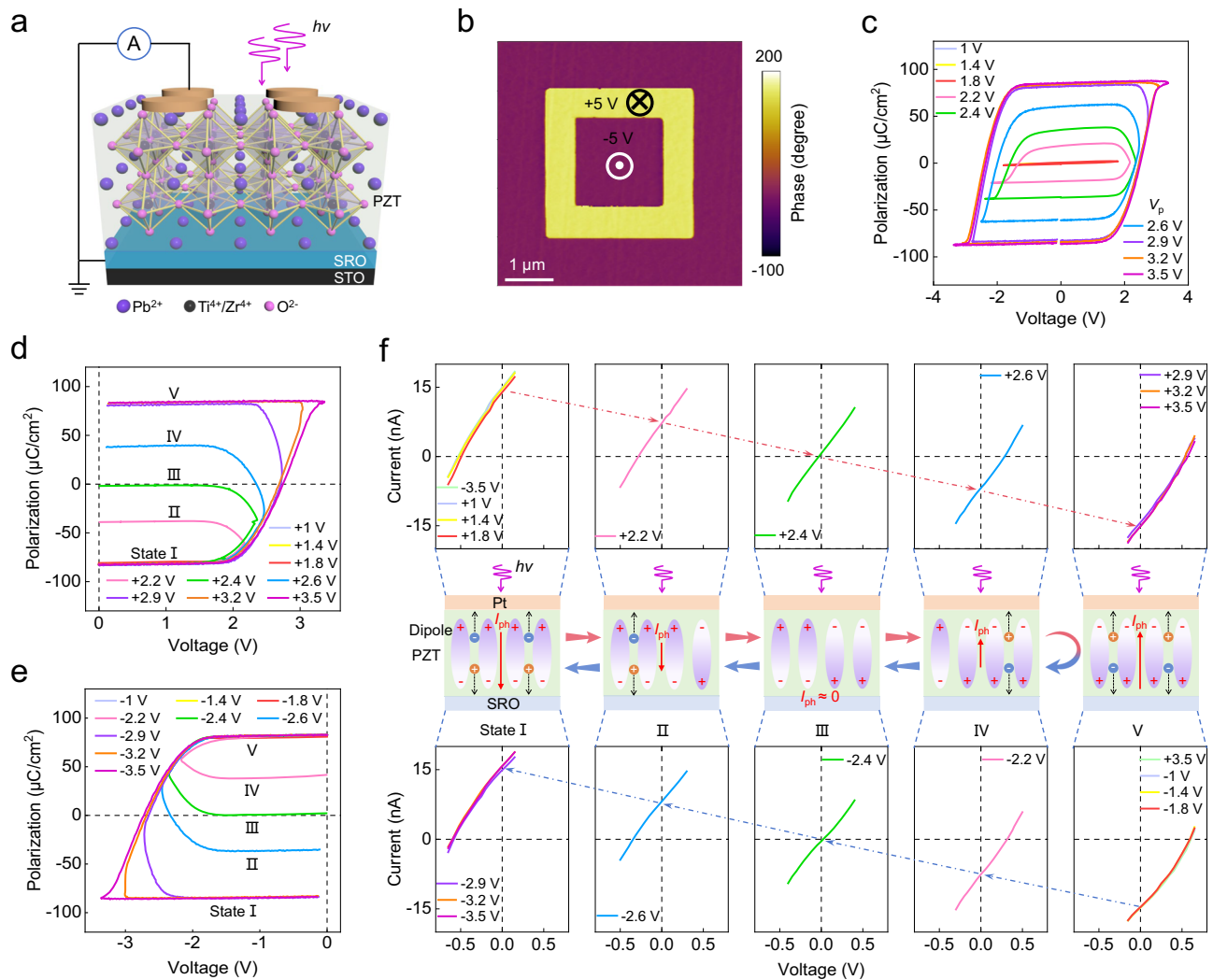
Nevertheless, the *in situ* training places high demands on the device performance. Specifically, nonvolatile multilevel weights, linear

and symmetric weight update, fast write and read speeds, high endurance, and small variations are typically required by the *in situ* training<sup>28</sup>. However, the existing programmable photosensors<sup>9,11</sup> struggle to satisfy all of these performance requirements. In addition, the *in situ* training also demands an efficient programming scheme for weight update<sup>29,30</sup>. While the programming schemes for in-memory ANNs have been extensively studied<sup>31–33</sup>, there is a lack of research on identifying efficient programming schemes for in-sensor ANNs. The demands for both high-performance devices and efficient programming schemes pose a significant challenge to the implementation of *in situ* training on in-sensor ANNs.

Here, we experimentally demonstrate the *in situ* training of an in-sensor ANN based on ferroelectric photosensors (FE-PSs) using a bi-directional closed-loop (BD-CL) programming scheme. The FE-PS is a programmable photosensor which operates through polarization control of photoresponsivity. It has been recently exploited as a building block of an in-sensor ANN, referred to as the ferroelectric



**Fig. 1 | Concepts of *ex situ* training and *in situ* training.** Schematics illustrating **a** *ex situ* training and **b** *in situ* training on in-sensor ANNs. This figure highlights the major advantages of the *in situ* training, including avoiding the use of additional digital computers, incorporating self-adaptive weight update, and being able to handle new data.



**Fig. 2 | Tunable polarization and photoresponse in FE-PS. a** Schematic of the device structure of the Pt/PZT/SRO FE-PS. **b** PFM phase image after writing a box-in-box pattern (outer: +5 V; inner: -5 V) on the bare PZT film.  $V_p$ -dependent **c** bipolar, **d** positive monopolar, and **e** negative monopolar  $P$ - $V$  hysteresis loops for the device. **f** Top and bottom panels: illuminated  $I$ - $V$  curves measured after applying positive and negative pulses with different  $V_p$ , respectively. Middle panel:

schematics of the different polarization states corresponding to the different  $V_p$ , which produce photocurrents with different magnitudes and directions. These schematics are drawn based on the results in (d, e). In (d-f) and anywhere else, the States I to IV refer to the full  $P_{up}$ , half  $P_{up}$ , near-zero-polarization, half  $P_{down}$ , and full  $P_{down}$  states, respectively.

photosensor network (FE-PS-NET), which demonstrates capability to implement real-time image sensing and recognition<sup>18–20</sup>. However, the in situ training of the FE-PS-NET remains unexplored hitherto, leaving uncertainty about whether the performance of the FE-PS is sufficient for the in situ training. We first show that the FE-PS with a device structure of Pt/Pb(Zr<sub>0.2</sub>Ti<sub>0.8</sub>)O<sub>3</sub> (PZT)/SrRuO<sub>3</sub> (SRO) exhibits symmetrically switchable photovoltaic responses as controlled by the remanent polarization. Besides, the FE-PS displays outstanding performance critical for in situ training, including self-powered, fast (<30  $\mu$ s), and multilevel (>4 bits) photoresponses, as well as long retention (50 days), high endurance (10<sup>9</sup>), high write speed (100 ns), and small cycle-to-cycle (C2C) and device-to-device (D2D) variations (-0.66% and -2.72%, respectively). Then, a BD-CL programming scheme is developed, achieving a precise and efficient weight update for the FE-PS. Based on the high-performance FE-PSs and the BD-CL programming scheme, we demonstrate that the FE-PS-NET is capable of being trained in situ to recognize traffic signs for commanding a prototype autonomous vehicle. Besides, the trained FE-PS-NET also exhibits high reliability (retaining 100% recognition accuracy for up to 50 days), high inference speed (50 times faster than a von Neumann

machine vision system), and zero energy consumption for inference (excluding contributions from peripheries). These results, therefore, showcase that the FE-PS-NET is a promising candidate for the development of in-sensor ANNs with in situ training capability.

## Results

### High-performance ferroelectric photosensors

Figure 2a schematically illustrates the device structure of our FE-PS, consisting of an epitaxial ferroelectric PZT film sandwiched between top Pt and bottom SRO electrodes. The detailed fabrication processes of the Pt/PZT/SRO FE-PS are presented in the “Methods” section. The Pt electrodes are -10 nm thick and -200  $\mu$ m in diameter. The PZT and SRO layers (thicknesses: -120 and -40 nm, respectively) are epitaxially grown on the STO substrate. The high epitaxial quality of the PZT/SRO film is revealed by the X-ray diffraction (XRD) results (Supplementary Fig. S1) and transmission electron microscopy (TEM) images (Supplementary Fig. S2). The atomic force microscopy (AFM) image (Supplementary Fig. S3a) shows a smooth surface of the PZT film with a small root-mean-square roughness of -440 pm, further confirming the good film quality. The piezoresponse force microscopy (PFM) imaging

(Fig. 2b) and hysteresis loops (Supplementary Fig. S3b) demonstrate the ferroelectricity of the PZT film. The high-quality ferroelectric PZT film provides a prerequisite for the manipulation of polarization and photoresponsivity in the Pt/PZT/SRO FE-PS.

The polarization switching behavior in the Pt/PZT/SRO FE-PS is characterized by measuring bipolar and monopolar polarization-voltage ( $P$ - $V$ ) hysteresis loops. In these measurements, the voltage is applied to the top Pt electrode while the bottom SRO electrode is grounded. The applied voltage pulses have a fixed width (0.15 ms) and varied amplitudes ( $V_p$ ). The measured  $V_p$ -dependent bipolar  $P$ - $V$  loops are shown in Fig. 2c. The loop begins to open when  $V_p$  reaches 2.2 V. As  $V_p$  continues to increase, the loop grows larger, eventually becoming nearly saturated at  $V_p = 2.9$  V. The saturated  $P$ - $V$  loop displays a remanent polarization ( $P_r$ ) as large as  $-80 \mu\text{C}/\text{cm}^2$ , consistent with the  $P_r$  values of high-quality epitaxial PZT films<sup>34–36</sup>. In addition, the saturated  $P$ - $V$  loop exhibits almost no voltage offset, suggesting that imprint field in our device is negligible. Because the imprint field is often defect-induced<sup>37–39</sup>, its absence in turn implies the high quality of the PZT film.

Figure 2c also demonstrates that multiple intermediate polarization states can be accessed by applying  $V_p$  in the range of 1.8–2.9 V. To further verify the accessibility of intermediate polarization states, monopolar  $P$ - $V$  loops are measured. Before measuring each monopolar  $P$ - $V$  loop, a full polarization up ( $P_{\text{up}}$ ) or down ( $P_{\text{down}}$ ) state is preset by applying a  $-3.5$  V or  $+3.5$  V pulse, respectively. As shown in Fig. 2d, the device remains in the full  $P_{\text{up}}$  state ( $-80 \mu\text{C}/\text{cm}^2$ ) when applying positive pulses with  $V_p \leq +1.8$  V. By increasing  $V_p$ , the device enters into 3 distinct intermediate states: half  $P_{\text{up}}$  ( $-40 \mu\text{C}/\text{cm}^2$ ), near-zero-polarization, and half  $P_{\text{down}}$  ( $+40 \mu\text{C}/\text{cm}^2$ ), at  $V_p = +2.2$ ,  $+2.4$ , and  $+2.6$  V, respectively. Further enhancing  $V_p$  to  $+2.9$  V and above results in the full  $P_{\text{down}}$  state ( $+80 \mu\text{C}/\text{cm}^2$ ). Conversely, applying negative pulses with increasing  $V_p$  can switch the device from the full  $P_{\text{down}}$  state to the full  $P_{\text{up}}$  state gradually (Fig. 2e). Notably, the up-to-down and down-to-up polarization switching processes are rather symmetric, benefiting from the aforementioned absence of imprint field. In addition, the loops in Fig. 2d, e feature relatively flat tops and bottoms, respectively, suggesting that the polarizations are nonvolatile when external voltages return to zero. The polarization nonvolatility is further verified by the retention test, where a long retention time of at least 24 h is observed (Supplementary Fig. S4).

The outstanding polarization switching characteristics exhibited by the Pt/PZT/SRO FE-PS, including large  $P_r$ , symmetric switching, multi-states, and nonvolatility, make it a promising candidate for investigating the polarization-modulated photovoltaic behavior. We first write the device into a specific polarization state by applying a  $-3.5$  V ( $+3.5$  V) preset pulse followed by a positive (negative) pulse  $V_p$  (the protocol is the same as that used in Fig. 2d,e). In each polarization state, the current-voltage ( $I$ - $V$ ) characteristics are measured under 365 nm ultraviolet (UV) illumination. Note that this light wavelength agrees well with the bandgap of PZT ( $-3.6$  eV)<sup>35,40,41</sup>. Hereafter unless otherwise specified, the applied light intensity and corresponding optical power (i.e., light intensity times electrode area) are  $-85 \text{ mW}/\text{cm}^2$  and  $-26.7 \mu\text{W}$ , respectively. As shown in the top panel of Fig. 2f, the device exhibits switchable photovoltaic responses apparently. More specifically, in the  $-3.5$  V-written state, the illuminated  $I$ - $V$  curve mainly locates in the second quadrant, displaying a short-circuit current ( $I_{\text{SC}}$ ) of  $-15 \text{ nA}$  and an open-circuit voltage ( $V_{\text{OC}}$ ) of  $-0.6 \text{ V}$ . This  $I_{\text{SC}}$  is 3 orders of magnitude higher than the dark current ( $-0.015 \text{ nA}$  @  $-0.6 \text{ V}$ ; see Supplementary Fig. S5). The illuminated  $I$ - $V$  curve shows little change after applying positive pulses with  $V_p \leq +1.8$  V. This is in accordance with the unchanged full  $P_{\text{up}}$  state at  $V_p \leq +1.8$  V (Fig. 2d). As  $V_p$  increases to  $+2.2$ ,  $+2.4$ ,  $+2.6$ , and  $+2.9$  V, the illuminated  $I$ - $V$  curve shifts from the second quadrant to the fourth quadrant. This is reminiscent of the polarization state evolution as shown in Fig. 2d. Further increasing  $V_p$  to  $+3.2$  V and above causes almost no shift of the illuminated  $I$ - $V$  curve.

This agrees with the saturation of the full  $P_{\text{down}}$  state at such large  $V_p$  (Fig. 2d). One can therefore establish a one-to-one correlation between the photoresponsivity states and the polarization states (see the top and middle panels of Fig. 2f).

In addition, applying negative pulses can switch the photoresponsivity states reversely (see the bottom panel of Fig. 2f), which correspond one-to-one to the polarization states from the full  $P_{\text{down}}$  state to the full  $P_{\text{up}}$  state (Fig. 2e). These results demonstrate that the photoresponse in our FE-PS is well controlled by the remanent polarization. The detailed mechanism underlying the polarization control of photoresponse is described in Supplementary Fig. S6. Notably, using the remanent polarization as the control knob of the photoresponse, the FE-PS is essentially a self-powered photosensor consuming zero energy for photosensing.

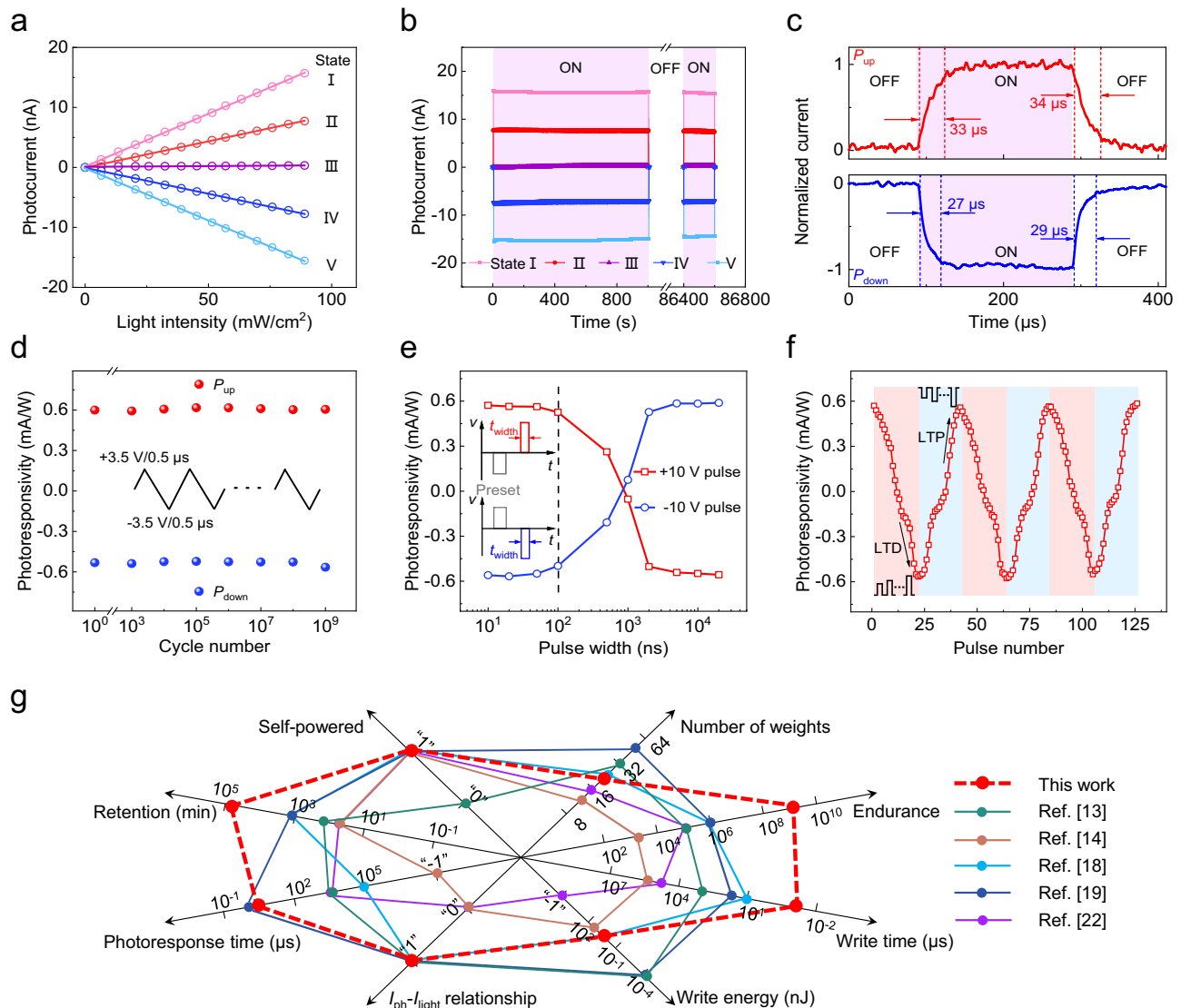
Another important feature of the FE-PS is the symmetric switching of photoresponse. As shown in Fig. 2f, the  $I_{\text{SC}}$  (or  $V_{\text{OC}}$ ) values in a pair of opposite polarization states have almost the same magnitude but opposite signs. For example, the  $I_{\text{SC}}$  value in the full  $P_{\text{up}}$  state is  $-15 \text{ nA}$ , which is just opposite to that in the full  $P_{\text{down}}$  state, i.e.,  $+15 \text{ nA}$ . Such symmetry is attributed to the fact that the control knob of the photoresponse, i.e., the polarization, exhibits symmetric switching owing to the absence of imprint field (Fig. 2c). Note that the symmetric switching of photoresponse is particularly useful, because it enables a single FE-PS to represent both positive and negative weights. Consequently, there is no need to use a pair of devices to represent a signed weight, greatly reducing the hardware overhead for network construction.

Next, the performance of the FE-PS as a programmable photosensor for in situ training is comprehensively investigated. In the in situ training, the FE-PS is mainly used for inference and weight update. We first focus on the performance metrics related to the inference. The inference accuracy is largely determined by the linearity of photocurrent versus light intensity. Figure 3a illustrates the measured relationships between photocurrent and light intensity for the 5 representative states, all of which can be well represented by linear fits (average coefficient of determination: 0.9997). This linear dependency of photocurrent on light intensity has been validated across different devices (Supplementary Fig. S7), allowing for precise multiplication between photoresponsivity and optical power. A high inference accuracy also requires a long retention of photoresponsivity. As shown in Fig. 3b, the photocurrents in different states are rather stable and can be retrieved after 24 h. Moreover, the photocurrent responses are highly reproducible in the cyclic test (Supplementary Fig. S8), and the photoresponsivity retention time can even be extended to 50 days (Supplementary Fig. S9). These results demonstrate the nonvolatility of the programmed photoresponsivities, which is well attributed to the polarization nonvolatility (Supplementary Fig. S4).

Besides the inference accuracy, the inference speed and energy consumption are also important concerns. A high inference speed demands a short photoresponse time. As seen from Fig. 3c, the average 10–90% photocurrent generation and decay times are both  $\sim 30 \mu\text{s}$ . In fact, these times are limited by the amplifying circuit used for measurement (Supplementary Fig. S10). Also considering that previous studies observed an ultrashort photoresponse time of  $<1 \text{ ns}$  in FE-PS<sup>42,43</sup>. We therefore infer that the photoresponse time of our FE-PS may be far below  $30 \mu\text{s}$ , thus allowing a fast inference speed. In addition, as mentioned earlier, our FE-PS works in the self-powered photovoltaic mode without external biases. Consequently, zero energy is consumed at the device level for the inference.

Then, we investigate the performance metrics related to the weight update. During the in situ training, the weight, i.e., photoresponsivity, needs to be frequently updated. Endurance is thus our first concern. The endurance test for the FE-PS is performed by applying cyclic  $\pm 3.5 \text{ V}/0.5 \mu\text{s}$  pulses. As displayed in Fig. 3d and Supplementary Figs. S11 and S12, both the photoresponsivity and





**Fig. 3 | Performance of the FE-PS. a** Light intensity dependence and **b** long-term stability of photocurrents of the device in the States I to IV. **c** Transient current responses to illumination for the device in the full  $P_{up}$  (upper panel) and  $P_{down}$  (lower panel) states. **d** Photoresponsivities of the device in the full  $P_{up}$  and  $P_{down}$  states after different endurance cycles. **e** Photoresponsivity as a function of write pulse width. The device is preset into a fully  $P_{up}$  ( $P_{down}$ ) state before each application of a +10 V (−10 V) write pulse with a varied width. **f** LTP/LTD characteristics

measured with an amplitude-increasing pulse scheme. **g** Performance comparison between our FE-PS and other emerging programmable photosensors for in-sensor computing. The “0” and “1” on the “self-powered” axis represents “not self-powered” and “self-powered”, respectively. The “0” and “1” on the “ $I_{ph}$ - $I_{light}$  relationship” represents “nonlinear” and “linear” relationships between photocurrent and light intensity, respectively. The “−1” on the “photoresponse time” and “write energy” axes refers to that the value is not reported.

polarization show little changes after  $10^9$  cycles. This is, to the best of knowledge, the highest endurance among the reported values for programmable photosensors<sup>19,44</sup>. Besides, the FE-PS also demonstrates a high speed for weight update (i.e., write speed), which is conducive to accelerating the training process. As illustrated in Fig. 3e, the photoresponsivity of the FE-PS can be tuned by  $\pm 10$  V pulses with widths as short as 100 ns. This write speed is at least one order of magnitude faster than those of other programmable photosensors<sup>19,20,44</sup>. However, the energy consumption for weight update (i.e., write energy) of our FE-PS is relatively high, reaching  $\sim 1.77$  nJ per operation (Supplementary Fig. S13). The relatively high write energy is mainly attributed to the large area of the present FE-PS ( $3.14 \times 10^4 \mu\text{m}^2$ ). To address this issue, device downscaling is a viable solution. For example, by scaling down the FE-PS to  $\sim 1 \mu\text{m}^2$  as demonstrated previously<sup>45</sup>, its write energy could be reduced to  $\sim 56$  fJ per operation, a sufficiently low value compared with those of other programmable photosensors<sup>14,17,19</sup>.

When performing the weight update in an open-loop manner, linear and symmetric modulation of multilevel photoresponsivities is preferred. To investigate it, long-term potentiation and depression (LTP and LTD, respectively) characteristics of the FE-PS are measured by employing an amplitude-increasing pulse scheme. Specifically, 21 positive pulses (amplitude: from 1.1 to 2.1 V in increments of 0.05 V; width: 3 ms) and 21 negative pulses (amplitude: from −1.1 to −2.1 V in decrements of −0.05 V; width: 3 ms) are applied successively to modulate the photoresponsivity. The use of a relatively large pulse width is due to the limitation of the pulse generator used in this measurement (see “Methods”). At this pulse width, the applied pulse voltages are around the coercive voltages (Supplementary Fig. S14), enabling them to switch the polarization and associated photoresponsivity. As shown in Fig. 3f, the photoresponsivity reduces from 0.56 to  $\sim 0.56$  mA/W with the positive pulses, and then increases back to 0.56 mA/W with the negative pulses. Such LTD and LTP processes are repeatable. Each LTP or LTD process contains 21 distinct photoresponsivity states,

confirming the FE-PS's capability to store multi-bit weights ( $>4$  bits). In fact, such a weight number is relatively moderate when compared to those reported previously<sup>14,19,22</sup>. To increase the weight number, a viable solution is to design a pulse scheme that enables a more gradual modulation of photoresponsivity, which warrants further investigation. In addition, the modulation of photoresponsivity is observed to be symmetric yet nonlinear, which is well associated with the symmetric, nonlinear polarization switching (Fig. 2c). The nonlinearity in photoresponsivity modulation can deviate a weight from its target value when updating the weight in an open-loop manner. To address this issue, one approach is to enhance the linearity of polarization switching through domain engineering<sup>46</sup>, a topic that warrants further exploration. Alternatively, a closed-loop programming scheme<sup>30,47</sup> can also be employed as a solution, as demonstrated later in this study. Besides, this scheme can also alleviate the adverse impacts of C2C and D2D variations on the precision of weight update, although the C2C and D2D variations of the FE-PS are quite small (only  $-0.66\%$  and  $-2.72\%$ , respectively; see Supplementary Fig. S15).

Figure 3g summarizes the performance of our FE-PS in comparison to other programmable photosensors for in-sensor computing, with a more detailed summary presented in Supplementary Table S1. It is shown that our FE-PS is one of the few self-powered devices showing simultaneously linear  $I_{ph}/I_{light}$  relationship and fast photoresponse. Furthermore, our FE-PS exhibits significantly longer retention, higher endurance, faster write speed, and smaller C2C and D2D variations compared to other programmable photosensors, particularly those used for in situ training<sup>9,11</sup>. Such superior performance of our FE-PS can mainly be attributed to its distinctive operation mechanism and high-quality epitaxial PZT film. While other programmable photosensors mainly rely on a volatile gating effect<sup>5,9–12,48</sup> or a kinetically slow defect migration effect<sup>13,49</sup> to tune photoresponsivity, our FE-PS operates through the polarization control of photoresponsivity. The photo-excited charge carriers in the FE-PS are separated by a polarization-induced asymmetric potential that is electrically switchable. This intrinsic process, without involving defect-mediated effects such as charge trapping/detrapping, allows a fast and self-powered photoresponse as well as a linear  $I_{ph}/I_{light}$  relationship. In addition, the high-quality epitaxial PZT film offers large nonvolatile polarization and excellent polarization switching properties for the FE-PS. As the polarization is the control knob of the photoresponsivity, our FE-PS could thus exhibit exceptionally long retention, high endurance, and fast write speed. These results demonstrate that our FE-PS is a competitive candidate for constructing an in-sensor ANN with in situ training capability.

### Programming schemes for weight update

Besides the performance of FE-PS, the programming scheme is also important for the implementation of in situ training. In the previous studies reporting the in situ training of in-sensor ANNs<sup>9,11</sup>, only an open-loop (OL) programming scheme was used for weight update. In contrast, this study aims to explore the optimal programming scheme, and hence 3 different programming schemes: OL, uni-directional closed-loop (UD-CL), and BD-CL, are comparatively investigated. Figure 4a shows the schematic of the OL scheme, where write pulses are applied to the device without verifying its photoresponsivity. However, both the BD-CL and UD-CL schemes apply a read operation after each write pulse to verify the device's photoresponsivity (see Fig. 4b, c, respectively). The main difference between the BD-CL and UD-CL schemes is the polarity of write pulses. In the UD-CL scheme, write pulses with the same polarity are applied to the device until its photoresponsivity reaches the target value within a certain margin of error. If the photoresponsivity exceeds the target value, known as being over-written, a refresh pulse is applied to re-initialize the photoresponsivity, followed by the repetition of unipolar write pulses (Fig. 4c). By contrast, in the BD-CL scheme, write

pulses with an opposite polarity are applied to correct any over-written photoresponsivity and bring it back towards the target value (Fig. 4b). The detailed flows of these programming schemes and their implementation platform are shown in Supplementary Figs. S16 and S17, respectively.

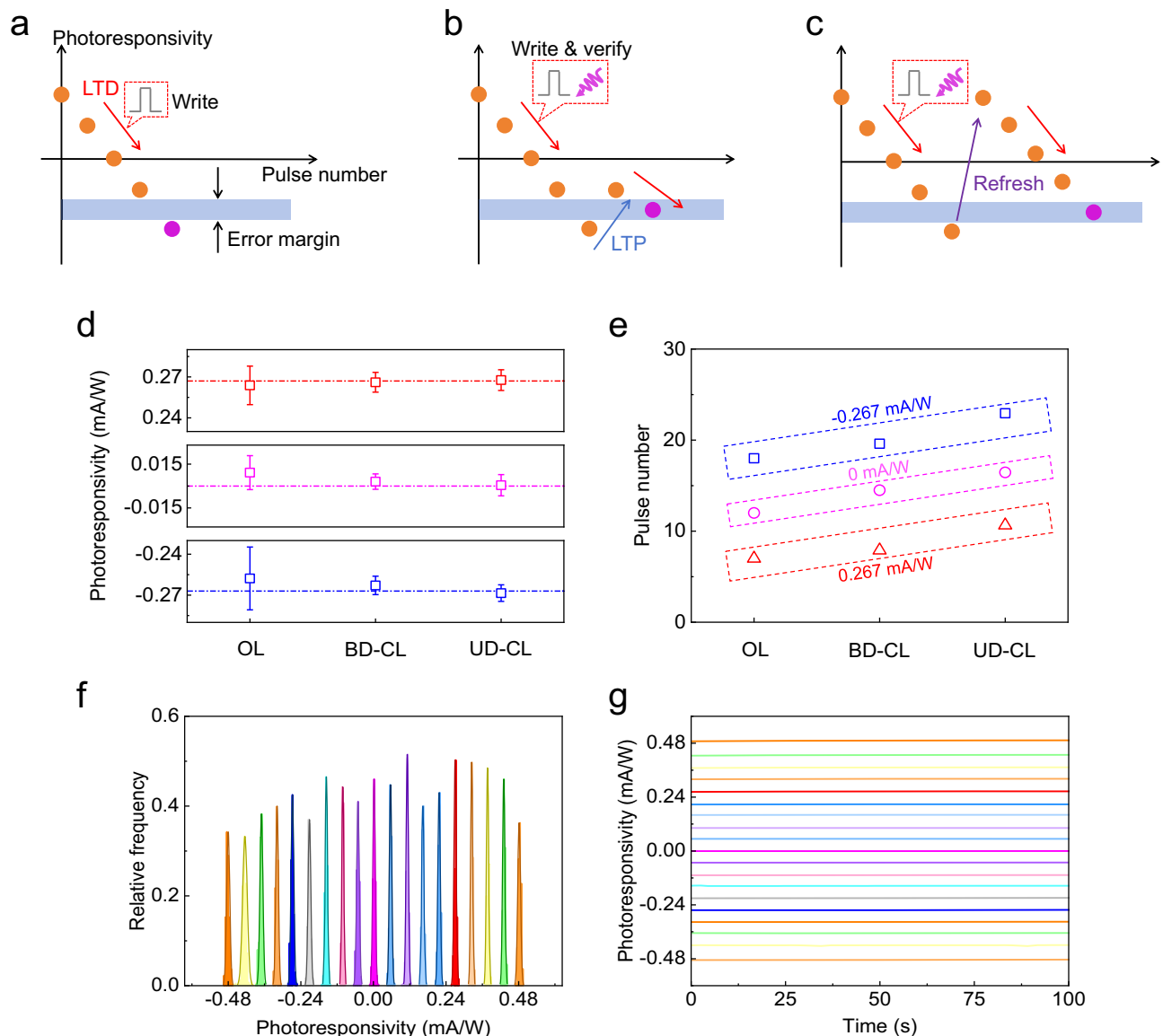
The above 3 programming schemes are compared in terms of programming precision and pulse consumption. A high programming precision is beneficial for achieving training convergence, while a low pulse consumption can help to save the time and energy costs in the training. For a fair comparison, the FE-PS is initialized at the same photoresponsivity of  $0.56$  mA/W. After the initialization, the device is programmed into 3 target photoresponsivities:  $0.267$ ,  $0$ , and  $-0.267$  mA/W, by using the 3 programming schemes. For each programming scheme, the initialization and programming are repeated for 20 times. Figure 4d shows that the UD-CL and BD-CL schemes achieve much lower discrepancies between actual and target photoresponsivities, i.e., higher programming precisions, compared to the OL scheme. This improvement is attributed to the use of verification with a small standard deviation margin ( $2.5\%$ ) in the CL schemes, which is absent in the OL scheme. This in turn confirms that the CL schemes can well address the nonlinearity in photoresponsivity modulation as well as C2C and D2D variations. On the other hand, as depicted in Fig. 4e, the pulse number increases as the programming scheme varies from OL to BD-CL and UD-CL. The lowest pulse consumption in the OL scheme is due to the absence of verification, and the combined use of verification and refresh causes the highest pulse consumption in the UD-CL scheme. These results demonstrate that the BD-CL scheme achieves the best tradeoff between programming precision and pulse consumption. In other words, the BD-CL scheme can realize a precise and efficient weight update for the FE-PS, and hence it is used hereafter for further study.

19 photoresponsivity states in the range of  $-0.48$  to  $0.48$  mA/W are repeatedly written for 20 times by using the BD-CL programming scheme. A standard deviation margin of  $2.5\%$  is used for the photoresponsivity verification. Figure 4f presents the distribution histogram of the 19 photoresponsivity states, which are well separated from each other without any overlap. In addition, these photoresponsivity states are demonstrated to be nonvolatile (Fig. 4g). These results further showcase the efficacy of the BD-CL programming scheme in attaining multi-bit weights in the FE-PS.

### In situ training of ferroelectric photosensor network

The availability of both the high-performance FE-PSs and the BD-CL programming scheme allows the construction of an FE-PS-NET with in situ training capability. To experimentally demonstrate it, 4 individual FE-PSs are connected in parallel to form a  $4 \times 1$  FE-PS-NET, which acts as a “visual system” of a prototype autonomous vehicle, as shown in Fig. 5a. The photographs of the FE-PS-NET and its peripheries in the vehicle are shown in Supplementary Fig. S18. The FE-PS-NET is trained in situ to learn the real-time recognition of traffic signs. The recognition result produced by the FE-PS-NET is directly sent to the motor system of the vehicle to control its movement. 4 traffic signs ( $2 \times 2$  pixels) are used for training, which represent the 4 commands: “go”, “stop”, “turn left”, and “turn right”, respectively (Fig. 5b). These traffic signs are temporarily implemented by a  $2 \times 2$  array of UV light-emitting diodes (LEDs), with one LED (i.e., one pixel) focusing on one FE-PS. The pixel values of “1” and “0” correspond to the optical powers of  $50$  and  $0$  mW/cm<sup>2</sup>, respectively.

Figure 5c depicts the flowchart of the in situ training, which mainly involves two processes: inference and weight update. In the inference process, a traffic sign is projected onto the FE-PS-NET. Each FE-PS in a pixel produces a photocurrent (or zero photocurrent) by multiplying its photoresponsivity by the optical power in this pixel, as shown in Fig. 5a. Simultaneously, the photocurrents of all the FE-PSs are



**Fig. 4 | Comparison of different programming schemes.** Schematics illustrating the modulation of photoresponsivity using **a** OL, **b** BD-CL, and **c** UD-CL programming schemes. **d** Measured photoresponsivities of the FE-PS after programming into 3 target photoresponsivities (from top to bottom: 0.267, 0, and -0.267 mA/W, respectively, as indicated by the dotted lines) using different programming schemes. **e** Numbers of pulses used to achieve the target photoresponsivities in **(d)**

for different programming schemes. In **(d, e)**, the reported photoresponsivity and pulse number values are averaged from 20 independent tests. The error bars in **(d)** indicate standard deviations. **f** Distribution histogram of 19 photoresponsivity states statistically obtained by repeatedly writing the device for 20 times (20 data points per one state per one time) using the BD-CL programming scheme. **g** Retention behavior of the 19 photoresponsivity states in **(f)**.

summed based on the Kirchhoff's law. The output current  $I_{\text{out}}$  is therefore given by

$$I_{\text{out}} = \sum_{n=1}^N R_n P_n \quad (1)$$

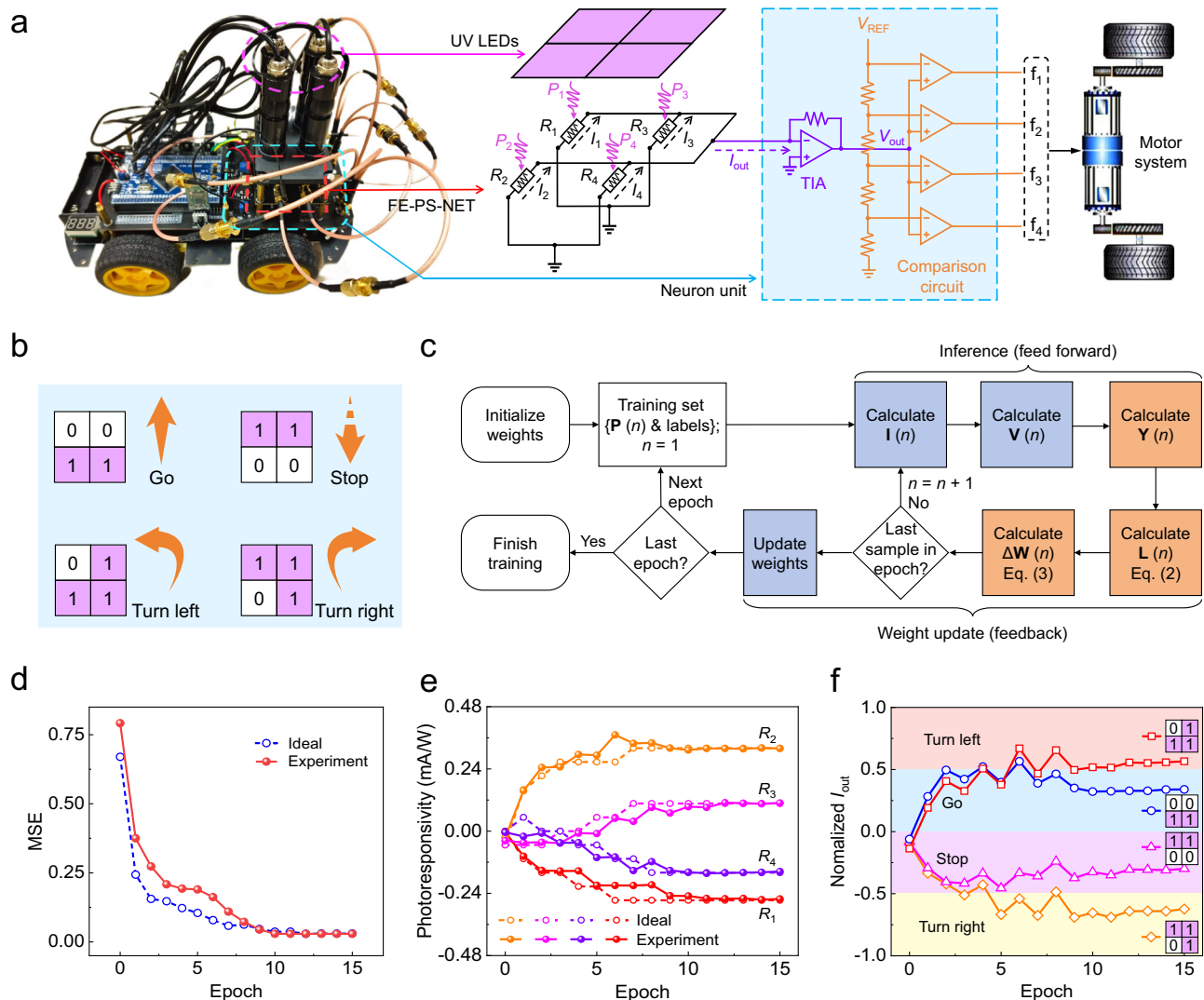
where  $R_n$  is the weight (i.e., photoresponsivity) at the  $n$ -th pixel,  $P_n$  is the input optical power at the  $n$ -th pixel, and  $N$  is the number of pixels. Equation (1) indicates that the FE-PS-NET can implement an in-sensor multiply-accumulate (MAC) operation, which is the basis for the real-time image sensing and recognition. The  $I_{\text{out}}$  is subsequently sent to a neuron unit comprising an amplifying circuit and a comparison circuit, as shown in Fig. 5a. The neuron unit converts the  $I_{\text{out}}$  to a voltage signal  $V_{\text{out}}$ , which is further represented by one of 4 levels: -0.75, -0.25, +0.25, and +0.75 V through voltage comparison. These 4 voltage levels

represent the recognition results of “turn right”, “stop”, “go”, and “turn left”, respectively.

The goal of the training is to minimize the discrepancy between the recognition results and the true labels of the traffic signs, which is measured by a mean square error (MSE) cost function implemented in software:

$$L = \frac{1}{M} \sum_{i=1}^M (y_i - \hat{y}_i)^2 \quad (2)$$

where  $y_i$  is the output after feeding the dimensionless value of the  $V_{\text{out}}$  for the  $i$ -th input image to a tanh activation function,  $\hat{y}_i$  is the true label of the  $i$ -th input image, and  $M$  is the number of input images. To minimize the cost function, the weights need to be updated to their optimal values. We use the gradient descent algorithm to guide the



**Fig. 5 | In situ training of an FE-PS-NET which acts as a “visual system” of a prototype autonomous vehicle. a** Photograph of the vehicle, along with schematic circuit diagrams. The output current of the FE-PS-NET is fed to the neuron unit, and then sent to the motor system of the vehicle. **b** Four traffic signs used for training and test. **c** Flowchart of the in situ training of the FE-PS-NET. The operations

in the light blue boxes are implemented in hardware, while those in the orange boxes are temporarily implemented in software on a PC and can be implemented by on-chip integrated circuits in the future. Evolutions of **d** MSE, **e** weights, and **f** normalized  $I_{out}$ s for different input traffic signs with the training epoch.

weight update. In this algorithm, the target weights are calculated in software based on the gradients of the cost function:

$$R_n := R_n - \alpha \frac{\partial L}{\partial R_n} \quad (3)$$

where the sign “:=” denotes assigning the value on the right side to the left side and  $\alpha$  is learning rate. Next, the photoreactivity of each FE-PS is experimentally adjusted to its corresponding target weight by using the BD-CL programming scheme. At this stage, one epoch of training is completed. With the platform shown in Supplementary Fig. S17, the training can proceed automatically for multiple epochs until the network converges.

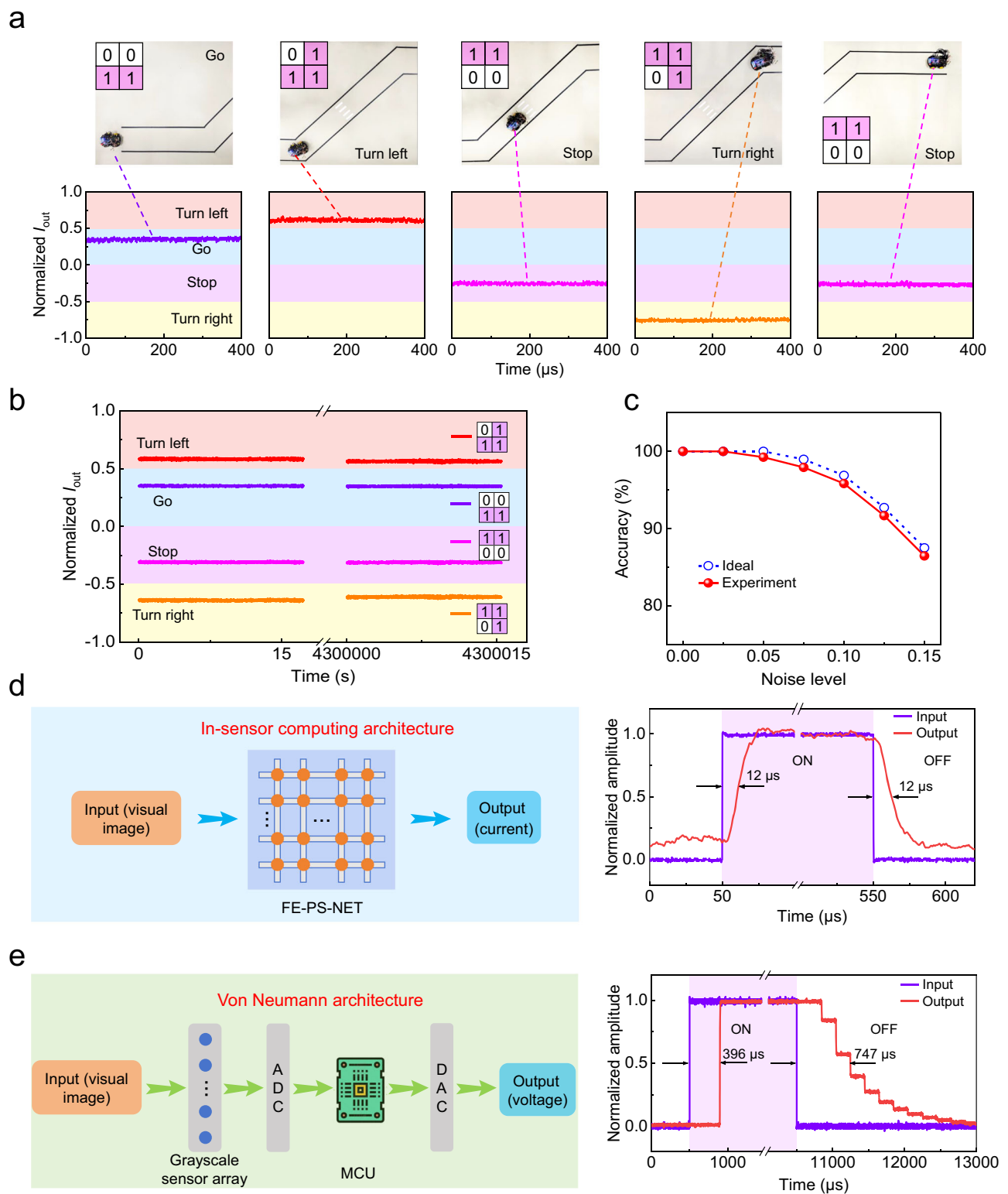
Figure 5d shows the evolution of the MSE during the in situ training of the FE-PS-NET. It is seen that the MSE decreases with the training epoch and becomes minimized after 10 epochs. The decrease in MSE is attributed to the smooth convergence of the weights, as presented in Fig. 5e. Notably, the training convergence behavior exhibited by the FE-PS-NET is well consistent with that of a software-based ANN (see comparison between the solid and dotted lines

in Fig. 5d, e), confirming the capability of the FE-PS-NET to implement the in situ training. Figure 5f displays the output currents of the FE-PS-NET for the 4 traffic signs at different training epochs. At Epoch #0, all the output currents fall into the range corresponding to the “stop” sign, indicating that the FE-PS-NET can only recognize the “stop” sign. As the training proceeds, more traffic signs are correctly recognized. The 100% recognition accuracy is achieved at Epoch #10.

The trained FE-PS-NET is then used for the test with the 4 traffic signs same as those used in the training. Figure 6a shows the navigation of the vehicle based on the real-time recognition results from the trained FE-PS-NET. The vehicle takes correct actions in response to all the traffic signs, well attributed to the accurate recognition provided by the trained FE-PS-NET. The behavior of the vehicle at the different training levels of the FE-PS-NET can be found in Supplementary Movie 1.

We further investigate the reliability of the trained FE-PS-NET. Figure 6b demonstrates that the trained FE-PS-NET retains 100% recognition accuracy after 50 days. Introducing noises into the test traffic signs (Supplementary Fig. S19) results in lower accuracies for the trained FE-PS-NET (Fig. 6c). However, it is worth noting that the





**Fig. 6 | Performance of the trained FE-PS-NET.** **a** Photographs of the vehicle when executing different motions (top) along with the corresponding  $I_{out}$ s of the FE-PS-NET (bottom). **b** Long-term stability of the  $I_{out}$ s of the FE-PS-NET. **c** Recognition accuracies of the FE-PS-NET at different noise levels (noise level refers to the

standard deviation of the Gaussian noise). **d** Schematic architecture of the FE-PS-NET (left) and its inference speed (right). **e** Schematic architecture of a von Neumann machine vision system (left) and its inference speed (right).

accuracy at each noise level remains close to its corresponding theoretical upper limit (see comparison between the solid and dotted lines in Fig. 6c). These results underscore the high reliability of the trained FE-PS-NET, which is well attributed to the nonvolatility of the programmed weights.

The weight nonvolatility also allows the FE-PS-NET to store the weights locally without the need for an external memory. Consequently, the FE-PS-NET can essentially integrate the sensing, memory, and processing functions. More importantly, it can function as an in-sensor ANN capable of performing high-level image processing with

boosted speed and energy efficiency. The inference speed of the FE-PS-NET is expected to be high as the image sensing and processing occurs simultaneously in the FE-PS-NET. Figure 6d shows the rising and falling delay times of the output current of the FE-PS-NET, both of which are as short as 12  $\mu$ s. The rising (falling) delay time is defined as the time lag between the 50% points of the rising (falling) edges of the input voltage (applied to the LED) and the output current. It should be noted that the measured delay times are indeed limited by the amplifying circuit used for measurement, similar to the issue encountered when measuring the photoresponse time of a single FE-PS. Nevertheless, the inference speed of the FE-PS-NET is still 50 times faster than that of a von Neumann machine vision system consisting of grayscale sensors and a microcontroller unit (MCU) (see comparison between Fig. 6d, e). Besides the high inference speed, the FE-PS-NET (excluding peripheries like the neuron unit) also exhibits zero energy consumption for inference because its constituent FE-PSs operate in the self-powered photovoltaic mode.

The speed and energy consumption of the FE-PS-NET for training are currently limited by the minimum pulse width (i.e., 3 ms) available in the in situ training experiment (see “Methods”). However, in the specific write speed and energy tests, the FE-PS demonstrates a fast write speed (100 ns; see Fig. 3e) and a potentially low write energy (~56 fJ per operation assuming device downscaling; see Supplementary Fig. S13), suggesting that the FE-PS-NET has the potential for achieving high training speed and energy efficiency. In addition, using on-chip integrated circuits to implement the operations which are temporarily realized by software in the in situ training process (Fig. 5c) could further boost the training speed and energy efficiency.

Although the FE-PS-NET demonstrated here is a small-scale network, it has the potential to be scaled up due to several factors. First, it has been demonstrated that the FE-PS could exhibit tunable photoresponsivity when downscaled to  $\sim 1 \mu\text{m}^2$ <sup>45</sup>. Further downscaling the FE-PS is viable because the commercial ferroelectric capacitor with the same device structure has been fabricated using the standard 130-nm complementary metal-oxide semiconductor (CMOS) process<sup>50</sup>. In addition, other merits of the FE-PS-NET, such as the local weight storage and the representation of both positive and negative weights in a single device, can substantially reduce the hardware overhead when scaling up the network. Therefore, the construction of a large-scale FE-PS-NET with a high area efficiency appears to be technologically feasible. Note that the large-scale FE-PS-NET empowered by the in situ training capability is competent for handling complicated machine vision tasks (see Supplementary Fig. S20 for demonstration). Besides the large-scale network construction, both the flexible design<sup>51</sup> of the FE-PS-NET and its integration with memristor-based neurons<sup>52</sup> are topics of great interest for future research.

## Discussion

In summary, we have experimentally demonstrated the in situ training of the FE-PS-NET using the BD-CL programming scheme. It is first shown that the building block of the FE-PS-NET, i.e., the Pt/PZT/SRO FE-PS, exhibits symmetrically switchable photovoltaic responses as controlled by the remanent polarization. Then, several key performance metrics related to both the inference and weight update in the in situ training are investigated for the FE-PS. In particular, the FE-PS displays self-powered, fast (<30  $\mu$ s) and multilevel (>4 bits) photoresponses, as well as linear dependency of photocurrent on light intensity, long retention (50 days), high endurance ( $10^9$ ), high write speed (100 ns), and small C2C and D2D variations (−0.66% and −2.72%, respectively). Next, several programming schemes are designed to implement the weight update in the FE-PS. Among them, the BD-CL programming scheme achieves the best tradeoff between programming precision and pulse consumption. Thanks to the high-performance FE-PSs and the BD-CL programming scheme, the in situ training of the FE-PS-NET to recognize traffic signs for commanding a

prototype autonomous vehicle is successfully implemented. Besides, the trained FE-PS-NET also shows high reliability (retaining 100% recognition accuracy for up to 50 days), high inference speed (50 times faster than a von Neumann machine vision system), and zero energy consumption for inference (excluding contributions from peripheries). Our study marks a significant advancement in the development of in-sensor computing systems with in situ training capability, which may be particularly useful for handling new data-streaming machine vision tasks.

## Methods

### Device fabrication

Approximately 120 nm PZT thin films together with ~40 nm SRO bottom electrodes were epitaxially grown on STO (001) substrates by pulsed laser deposition (PLD). The PLD system mainly comprised a KrF excimer laser source ( $\lambda = 248$  nm) (Coherent COMPexPro 205) and a customized chamber. Different laser energy fluences of 0.90 J/cm<sup>2</sup> and 0.97 J/cm<sup>2</sup> were used for the deposition of PZT and SRO films, respectively, while the repetition rates were kept the same at 5 Hz. The SRO films were first deposited at a substrate temperature of 680 °C under an oxygen pressure of 15 Pa, followed by the deposition of PZT films at a lower substrate temperature of 620 °C under the same oxygen pressure. Subsequently, the PZT/SRO films were cooled to room temperature at a 10 °C/min cooling rate under 1000 Pa oxygen pressure. Then, Pt top electrodes (thickness: ~10 nm; diameter: ~200  $\mu$ m) were ex situ deposited on the PZT/SRO films through a shadow mask by sputtering at room temperature and under vacuum. The resulting Pt/PZT/SRO capacitor-like heterostructures are the desired FE-PSs. To construct a FE-PS-NET, multiple FE-PSs were connected to a test board containing pre-fabricated interconnections between different cells (Supplementary Fig. S17).

### Characterizations

The phases and crystalline structures of the fabricated films were characterized by XRD (“X” Pert PRO, PANalytical). The microstructures were further investigated using TEM (Tecnai G2-F20). AFM and PFM studies were carried out on an integrated scanning probe microscope (Asylum Research MFP-3D) with Pt-coated silicon tips (Nanoworld EFM Arrow). The PFM images and hysteresis loops were measured by using an AC driving voltage of 0.8 V in the dual a.c. resonance tracking (DART) mode.

### Electrical measurements

Ferroelectric  $P$ – $V$  hysteresis loops were recorded using a ferroelectric workstation (Radiant Precision Multiferroic). DC  $I$ – $V$  characteristics and low-speed photoresponses were measured with a Keithley 6430 SourceMeter. High-speed photoresponses were measured by using a combination of an amplifying circuit and an oscilloscope (LeCory 64Xi-A). In all the photoresponse measurements, optical inputs were supplied by 365 nm UV LEDs with tunable light intensities. Voltage pulses were applied to tune the polarization and associated photoresponsivity; however, different pulse generators were used in different experiments. When performing the LTP/LTD measurement and implementing various programming schemes, the pulses (minimum width: 3 ms) were provided by 12-bit digital-to-analog converters (DACs) as controlled by an STM32 MCU. In the measurements of endurance and write speed, a function generator (Agilent 33250 A) was used to generate the pulses (minimum width: 10 ns). In the rest cases, the ferroelectric workstation was used as the pulse generator.

The setup for the in situ training of the FE-PS-NET is shown in Supplementary Fig. S17. It mainly included a test board carrying the FE-PS-NET, an amplifying circuit, an STM32 MCU, a 8-channel 16-bit analog-to-digital converter (ADC), 12-bit DACs, and a personal computer (PC). When performing the training, as shown in Fig. 5c, the FE-PS-NET produced a photocurrent upon the application of illumination. Subsequently, the current signal was sent to the amplifying circuit for

current–voltage conversion and amplification. The resulting voltage signal was further directed to the ADC, the MCU, and ultimately the PC. The PC implemented the gradient descent algorithm, and relayed the result back to the MCU. Communication between the PC and MCU was conducted via a universal asynchronous receiver/transmitter (UART). The MCU then instructed the DACs to apply pulses to the FE-PS-NET to update the weights. To realize a precise weight update, closed-loop programming schemes were executed (Supplementary Fig. S16). The training proceeded automatically for multiple epochs until the network became converged.

The FE-PS-NET trained at different levels was used to perform the real-time image recognition for a prototype autonomous vehicle. Besides the FE-PS-NET, the vehicle also comprised a motor system, a STM32 MCU, 2 batteries, and a vehicle body (Supplementary Fig. S18). The real-time recognition result of the FE-PS-NET was directly sent to the motor system of the vehicle to control its movements, while the MCU was used only for controlling LEDs.

## Data availability

The data that support the findings of this study are available in the article and the Supplementary Information. Additional data related to this study can be requested from the corresponding authors. Source data are provided with this paper.

## Code availability

All the codes that support the findings of this study are available from the corresponding authors upon request.

## References

- Feng, G., Zhang, X., Tian, B. & Duan, C. Retinomorphic hardware for in-sensor computing. *InfoMat* **5**, e12473 (2023).
- Wan, T. et al. In-sensor computing: materials, devices, and integration technologies. *Adv. Mater.* **35**, 2203830 (2023).
- Zhou, F. et al. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat. Nanotechnol.* **14**, 776–782 (2019).
- Chen, J. et al. Optoelectronic graded neurons for bioinspired in-sensor motion perception. *Nat. Nanotechnol.* **18**, 882–888 (2023).
- Huang, P. et al. Neuro-inspired optical sensor array for high-accuracy static image recognition and dynamic trace extraction. *Nat. Commun.* **14**, 6736 (2023).
- Li, G. et al. Photo-induced non-volatile VO<sub>2</sub> phase transition for neuromorphic ultraviolet sensors. *Nat. Commun.* **13**, 1729 (2022).
- Zhou, G. et al. Full hardware implementation of neuromorphic visual system based on multimodal optoelectronic resistive memory arrays for versatile image processing. *Nat. Commun.* **14**, 8489 (2023).
- Wu, X. et al. Wearable in-sensor reservoir computing using optoelectronic polymers with through-space charge-transport characteristics for multi-task learning. *Nat. Commun.* **14**, 468 (2023).
- Mennel, L. et al. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62–66 (2020).
- Wang, C. et al. Gate-tunable van der Waals heterostructure for reconfigurable neural network vision sensor. *Sci. Adv.* **6**, eaba6173 (2020).
- Dang, B. et al. One-phototransistor–one-memristor array with high-linearity light-tunable weight for optic neuromorphic computing. *Adv. Mater.* **35**, 2204844 (2023).
- Pi, L. et al. Broadband convolutional processing using band-alignment-tunable heterostructures. *Nat. Electron.* **5**, 248–254 (2022).
- Zhang, G. et al. Broadband sensory networks with locally stored responsivities for neuromorphic machine vision. *Sci. Adv.* **9**, eadi5104 (2023).
- Cai, Y. et al. Molecular ferroelectric/semiconductor interfacial memristors for artificial synapses. *npj Flex. Electron.* **6**, 16 (2022).
- Lee, S., Peng, R., Wu, C. & Li, M. Programmable black phosphorus image sensor for broadband optoelectronic edge computing. *Nat. Commun.* **13**, 1485 (2022).
- Kang, Y. et al. Bioinspired activation of silent synapses in layered materials for extensible neuromorphic computing. *J. Materiomics* **9**, 787–797 (2023).
- Cai, Y. et al. In-situ artificial retina with all-in-one reconfigurable photomemristor networks. *npj Flex. Electron.* **7**, 29 (2023).
- Cui, B. et al. Ferroelectric photosensor network: an advanced hardware solution to real-time machine vision. *Nat. Commun.* **13**, 1707 (2022).
- Wu, G. et al. Ferroelectric-defined reconfigurable homojunctions for in-memory sensing and computing. *Nat. Mater.* **22**, 1499–1506 (2023).
- Sun, Y. et al. Mesoscopic sliding ferroelectricity enabled photo-voltaic random access memory for material-level artificial vision system. *Nat. Commun.* **13**, 5391 (2022).
- Jang, H. et al. In-sensor optoelectronic computing using electrostatically doped silicon. *Nat. Electron.* **5**, 519–525 (2022).
- Wang, Y. et al. A three-dimensional neuromorphic photosensor array for nonvolatile in-sensor computing. *Nano Lett.* **23**, 4524–4532 (2023).
- Yao, P. et al. Face classification using electronic synapses. *Nat. Commun.* **8**, 15199 (2017).
- Li, C. et al. Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nat. Commun.* **9**, 2385 (2018).
- Wang, Z. et al. In situ training of feed-forward and recurrent convolutional memristor networks. *Nat. Mach. Intell.* **1**, 434–442 (2019).
- Ning, H. et al. An in-memory computing architecture based on a duplex two-dimensional material structure for in situ machine learning. *Nat. Nanotechnol.* **18**, 493–500 (2023).
- Shi, J. & Zeng, Z. Design of in-situ learning bidirectional associative memory neural network circuit with memristor synapse. *IEEE Trans. Emerging Top. Comput. Intell.* **5**, 743–754 (2021).
- Zhang, W. et al. Neuro-inspired computing chips. *Nat. Electron.* **3**, 371–382 (2020).
- Wang, I., Chang, C., Chiu, L., Chou, T. & Hou, T. 3D Ta/TaO<sub>x</sub>/TiO<sub>2</sub>/Ti synaptic array and linearity tuning of weight update for hardware neural network applications. *Nanotechnology* **27**, 365204 (2016).
- Li, C., An, J., Kweon, J. & Song, Y. Alternate pulse scheme in a hardware neural network for reducing the influence of asymmetry on synaptic weight updating. *J. Semicond. Tech. Sci.* **20**, 119–126 (2020).
- Mackin, C. et al. Optimised weight programming for analogue memory-based deep neural networks. *Nat. Commun.* **13**, 3765 (2022).
- Haensch, W. et al. Compute in-memory with non-volatile elements for neural networks: a review from a co-design perspective. *Adv. Mater.* **35**, 2204944 (2023).
- Li, N. et al. Optimization of projected phase change memory for analog in-memory computing inference. *Adv. Electron. Mater.* **9**, 2201190 (2023).
- Vrejoiu, I. et al. Intrinsic ferroelectric properties of strained tetragonal Pb<sub>0.2</sub>Ti<sub>0.8</sub>O<sub>3</sub> obtained on layer-by-layer grown, defect-free single-crystalline films. *Adv. Mater.* **18**, 1657–1661 (2006).
- Tan, Z. et al. Thinning ferroelectric films for high-efficiency photo-voltaics based on the Schottky barrier effect. *NPG Asia Mater.* **11**, 20 (2019).
- Lu, L. et al. Effects of interfacial buffering layer on imprint and domain switching dynamics in Pb(Zr,Ti)O<sub>3</sub> thin-film heterostructures. *J. Adv. Dielectr.* **14**, 2340010 (2024).
- Yu, J. et al. The effect of seeding layers on ferroelectric properties of PTZT thin films. *Integr. Ferroelectr.* **85**, 59–66 (2006).
- Pintilie, L., Stancu, V., Vasile, E. & Pintilie, I. About the complex relation between short-circuit photocurrent, imprint and polarization in ferroelectric thin films. *J. Appl. Phys.* **107**, 114111 (2010).

39. Buragohain, P. et al. Fluid imprint and inertial switching in ferroelectric La:HfO<sub>2</sub> capacitors. *ACS Appl. Mater. Interfaces* **11**, 35115–35121 (2019).
40. Gupta, R., Gupta, V. & Tomar, M. Ferroelectric PZT thin films for photovoltaic application. *Mater. Sci. Semicond. Process.* **105**, 104723 (2020).
41. Tan, Z. et al. Polarization imprint effects on the photovoltaic effect in Pb(Zr,Ti)O<sub>3</sub> thin films. *Appl. Phys. Lett.* **112**, 152905 (2018).
42. Xing, J. et al. High-sensitive switchable photodetector based on BiFeO<sub>3</sub> film with in-plane polarization. *Appl. Phys. Lett.* **106**, 033504 (2015).
43. Li, J. et al. Self-driven visible-blind photodetector based on ferroelectric perovskite oxides. *Appl. Phys. Lett.* **110**, 142901 (2017).
44. Kumar, D. et al. Artificial visual perception neural system using a solution-processable MoS<sub>2</sub>-based in-memory light sensor. *Light Sci. Appl.* **12**, 109 (2023).
45. Fan, H. et al. Large electroresistance and tunable photovoltaic properties of ferroelectric nanoscale capacitors based on ultrathin super-tetragonal BiFeO<sub>3</sub> films. *J. Mater. Chem. C* **5**, 3323–3329 (2017).
46. Luo, Z. et al. High-precision and linear weight updates by sub-nanosecond pulses in ferroelectric tunnel junction for neuro-inspired computing. *Nat. Commun.* **13**, 699 (2022).
47. Yi, W. et al. Feedback write scheme for memristive switching devices. *Appl. Phys. A* **102**, 973–982 (2011).
48. Zhang, T. et al. High performance artificial visual perception and recognition with a plasmon-enhanced 2D material neural network. *Nat. Commun.* **15**, 2471 (2024).
49. Chen, Q. et al. Switchable perovskite photovoltaic sensors for bioinspired adaptive machine vision. *Adv. Intell. Syst.* **2**, 2000122 (2020).
50. McAdams, H. P. et al. A 64-Mb embedded FRAM utilizing a 130-nm 5LM Cu/FSG logic process. *IEEE J. Solid-State Circuits* **39**, 667–677 (2004).
51. Lee, H. E. et al. Trichogenic photostimulation using monolithic flexible vertical AlGaInP light-emitting diodes. *ACS Nano* **12**, 9587–9595 (2018).
52. Sung, S. H., Kim, T. J., Shin, H., Im, T. H. & Lee, K. J. Simultaneous emulation of synaptic and intrinsic plasticity using a memristive synapse. *Nat. Commun.* **13**, 2811 (2022).

## Acknowledgements

The authors would like to thank the National Key Research and Development Programs of China (Grant No. 2022YFB3807603), the National Natural Science Foundation of China (Grant Nos. 92163210 and 52172143), the Science and Technology Projects in Guangzhou (Grant Nos. 202201000008 and 2022A04J00031), and the Guangdong Natural

Science Funds for Distinguished Young Scholar (Grant No. 2024B1515020053).

## Author contributions

Z.F. conceived the research. Z.F. and X.Y. supervised the project. H.L., J.O., and B.C. prepared the devices. H.L., J.O., W.H., W.L., G.Z., and X.G. performed the XRD, TEM, and PFM characterizations. H.L., J.O., J.X., Z.C., B.Y., K.L., L.M., M.L., X.L., and G.Z. conducted the electrical measurements. H.L. and J.O. carried out the simulations. H.L., J.O., Z.F., X.Y., X.G., and J.-M.L. wrote and revised the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-55508-z>.

**Correspondence** and requests for materials should be addressed to Zhen Fan or Xiaobing Yan.

**Peer review information** *Nature Communications* thanks Ming Wang, 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024