Supplementary Information

Birdlike broadband neuromorphic visual sensor arrays for fusion imaging

Pengshan Xie[1, #], Yunchao Xu[2,3 #], Jingwen Wang[2,3 #], Dengji Li[1], Yuxuan Zhang[1], Zixin Zeng[1], Boxiang Gao[1], Quan Quan[1], Bowen Li[1], You Meng[1,4], Weijun Wang[1], Yezhan Li[1], Yan Yan[1], Shen Yi[1], Jia Sun[2,3 *], Johnny C. Ho[1,4,5 *]

1 Department of Materials Science and Engineering, City University of Hong Kong, Hong Kong SAR 999077, China

2 Hunan Key Laboratory for Super-microstructure and Ultrafast Process, School of Physics, Central South University, Changsha 410083, Hunan, China.

3 State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, 865 Changning Road, Shanghai 200050, China

4 State Key Laboratory of Terahertz and Millimeter Waves, City University of Hong Kong, Hong Kong SAR 999077, China

5 Institute for Materials Chemistry and Engineering, Kyushu University, Fukuoka 816-8580, Japan

# These authors contributed equally: Pengshan Xie, Yunchao Xu, Jingwen Wang

* Corresponding authors E-mail: jiasun@csu.edu.cn (J. Sun), johnnyho@cityu.edu.hk (J. C. Ho)

Table of Contents

1. Simulations of the Reservoir Computing

1.1 Workflow of the Reservoir in Extracting Features of 2D Images: First, each pixel of the optical image is processed. The pixel value [0-255,0-255,0-255] is converted from RGB$_{888}$ format to RGB$_{565}$ format, being 6 bits long. This means that each time, an input step is either "1" (0.5-sec illumination) or "0" (0.5-sec darkness). During the simulation, the reservoir used for color recognition takes three 6-long optical sequences as input and produces output features based on the dynamic changes resulting from the input. The simulated crossbar array of memristor transistors can further classify the resulting three long features.

1.2 Workflow of the Reservoir in Extracting Features of Videos: The event video stream was first down-sampled to 200 pixels in spatial (X and Y) directions and then temporally subdivided into 84 frames (i.e., integrating events over 84 equally spaced intervals). Four evenly spaced frames were extracted from the 84 frames for downstream categorization. Like the feature extraction for the 2D images, each synaptic device pixel converts the 4 frames of optical event pulse flow into changes in source leakage current, thresholds these changes, and extracts the values of the current changes from the total matrix to visualize the numerical effects.

1.3. Supplementary Method 1: The color corresponding to RGB$_{888}$ format is converted into 3 wavelengths of light pulse strings according to the following rules. Each 0 or 1 represents if there is light or not. The conversion rules are as follows:

$$R_{565} = concatenate([0, R_{888}[:-3]]) \tag{1}$$

$$G_{565} = G_{888}[:-2] \tag{2}$$

$$B_{565} = concatenate([0, B_{888}[:-3]]) \tag{3}$$

$$UV_6 = concatenate([0, UV_8[:-3]]) \tag{4}$$

Taking R_{888} as an example, $R_{888}[:-3]$ stands for discarding the next 3 bits and taking only the first 5 bits, spliced with 1 bit of 0 to synthesize a 6-bit burst. For the UV

band, the luminance mapping is similar to RGB. Then, pulses are input device reservoirs that generate current in a final state ((3) and (4) in Supplementary Method 2).

1.4. Supplementary Method 2: Set the luminance in two neighboring frames as f_1 and f_2 respectively:
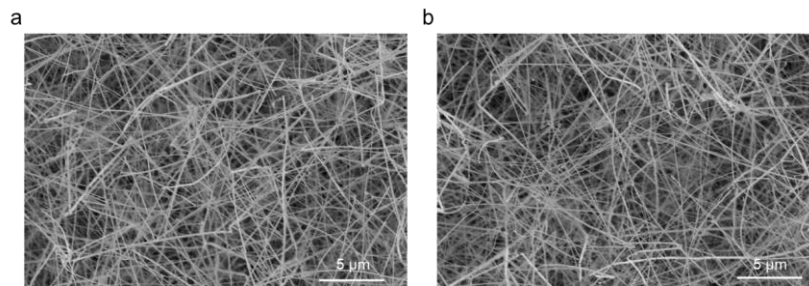
$$f_1 = +bin(f_1)[:-2] \tag{1}$$

$$f_2 = -bin(f_2)[:-2] \tag{2}$$

$$I_{out} = g([f_1, f_2]) \tag{3}$$

$$g(f) = device(f(t)) \tag{4}$$

$bin()$ represents the binary operation on the decimal value, $device()$ represents the dynamic response of the device to the burst, $f(t)$ and $g(f)$ represent respectively the input optical pulses and the resulting output of the input stream at the current time step.

2. Supporting Figures

Supplementary Fig. 1| Scanning electron microscopy (SEM) images of the GaAs on the Si/SiO$_2$ substrate with different regions. **a,** Closer to the source. **b,** Back end of growth substrate.
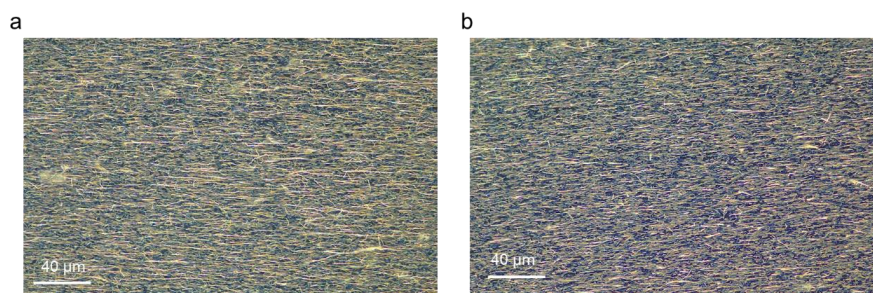
Supplementary Fig. 2| Optical images of the large-scale NW arrays fabricated by dry transfer technique. **a**, NWs arrays located at the front of the substrate. **b**, NWs arrays located at the back end of the substrate.
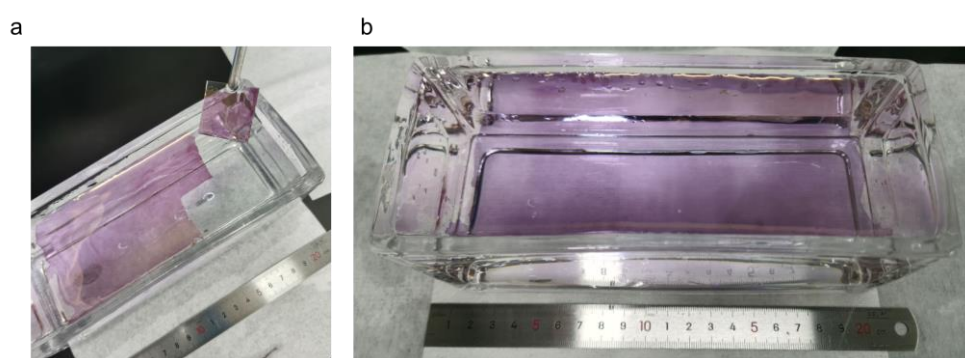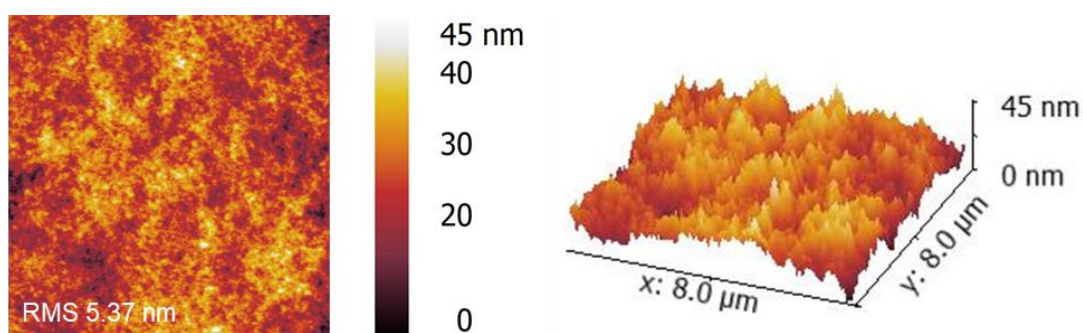

Supplementary Fig. 3| Self-assembly of the P3HT film on the water surface. **a**, Physical transfer of the film onto a glass substrate. **b**, Large area self-assembly P3HT film (on 20 cm × 8 cm water surface).
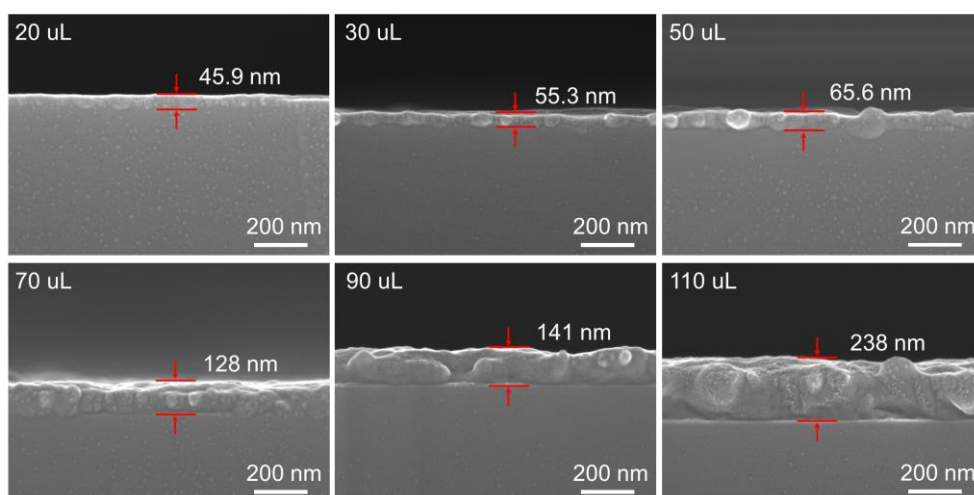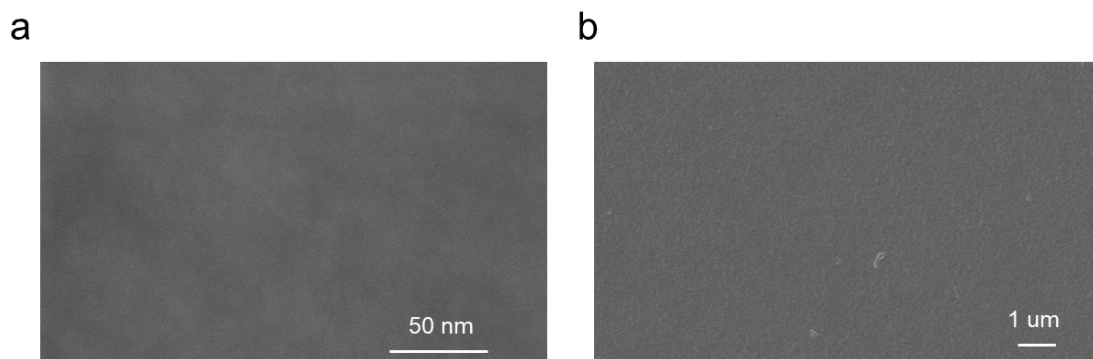

Supplementary Fig. 4| Cross-sectional SEM images with different self-assembled P3HT film thicknesses.
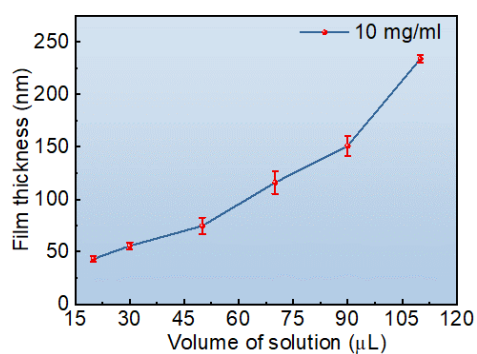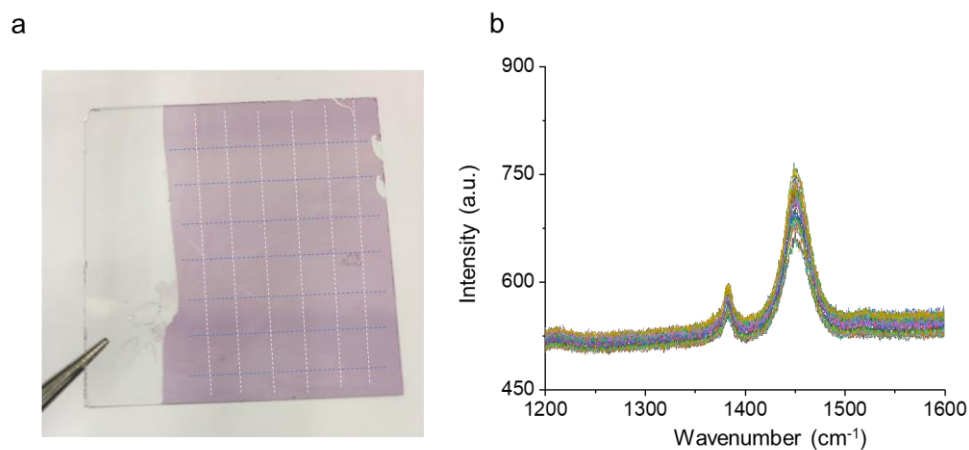
**Supplementary Fig. 5|** Corresponding volume-thickness relationship at the same concentration.

Supplementary Fig. 6| AFM images of the self-assembled P3HT film (10 mg/mL, 50 μL).
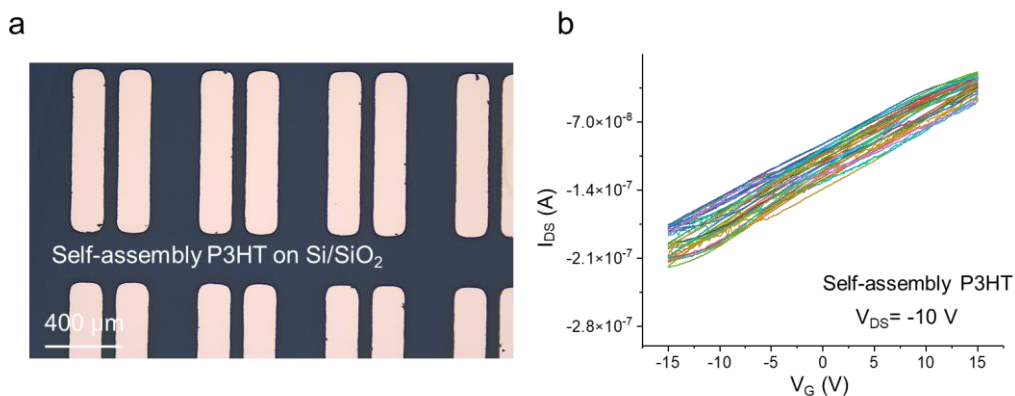

Supplementary Fig. 7| SEM images of the self-assembled P3HT film with different scale bar. **a**, 50 nm. **b**, 1 μm.

Supplementary Fig. 8| a, Photograph of the P3HT film self-assembled on glass. **b**, Raman spectrum of the randomly selected 42 locations on the film.
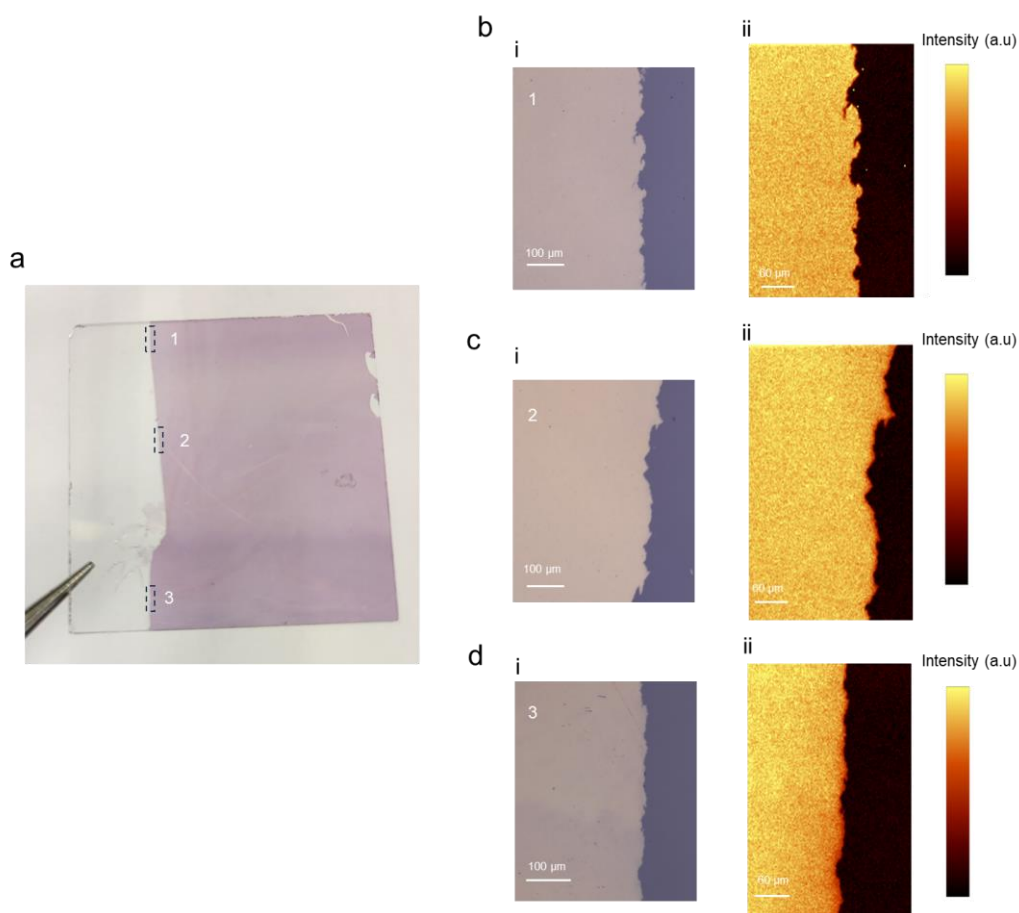

Supplementary Fig. 9| a, Photograph of the P3HT film self-assembled on glass. **b-d**, Raman spectrum mapping of the 3 areas selected on the film.
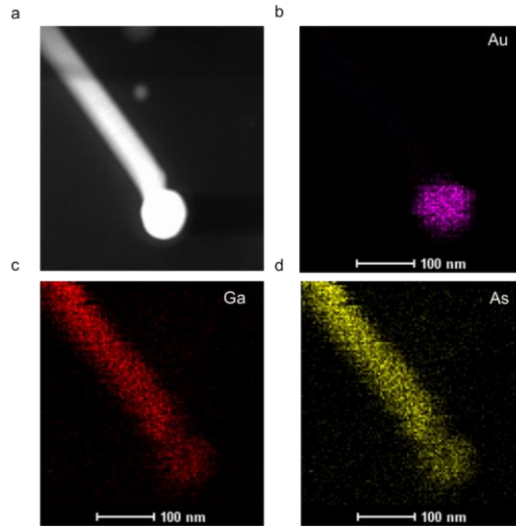
Supplementary Fig. 10| a, The self-assembled P3HT film on Si/SiO₂ with Ag S/D electrodes. **b**, Transfer characteristic curves of the randomly selected 28 devices.
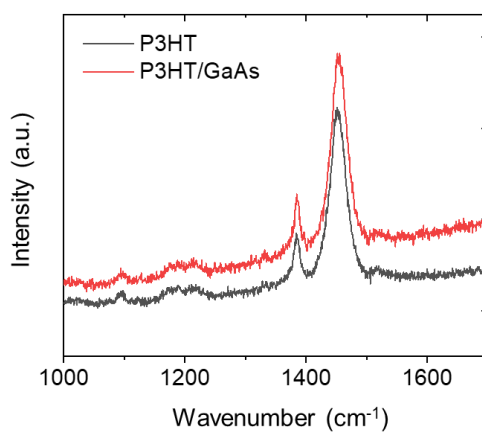

Supplementary Fig. 11| Raman spectra of the self-assembled P3HT film with and without GaAs NWs beneath it.
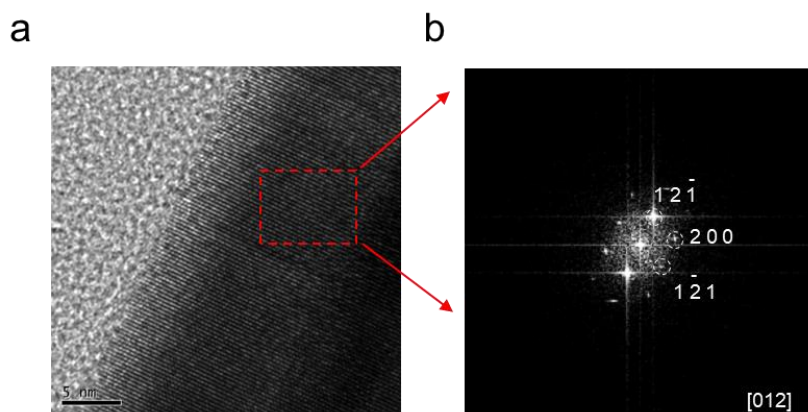
**Supplementary Fig. 12| a**, HRTEM image of the GaAs NW and **b**, the corresponding fast Fourier transform (FFT).

Supplementary Fig. 13| a, TEM image of single GaAs NW. EDS mapping of the GaAs NW with **b**, Au, **c**, Ga, **d**, As elements. The aggregation of the Au catalyst at the tip demonstrates its vapor-liquid-solid (VLS) growth process.
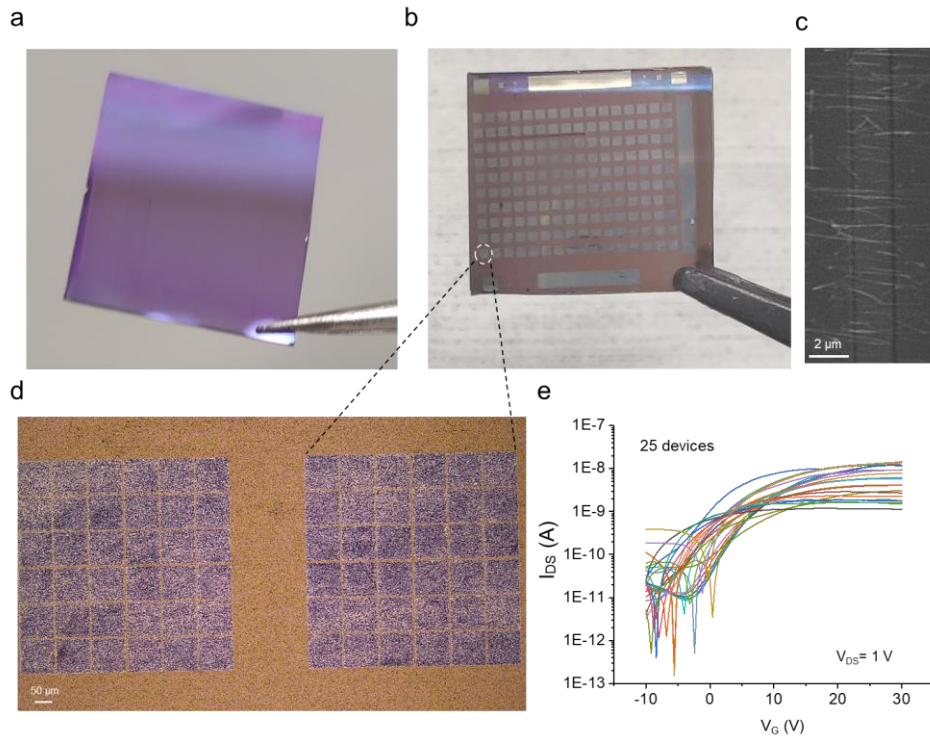

Supplementary Fig. 14| a, Dry contact transfer of NWs on the Si/SiO$_2$ substrate (n^{++} Si pieces pre-grown with the 200 nm thick SiO$_2$ dielectric layer). **b**, Photograph of the field-effect transistor arrays based on NWs. **c**, SEM image of the 2 µm channel. **d**,
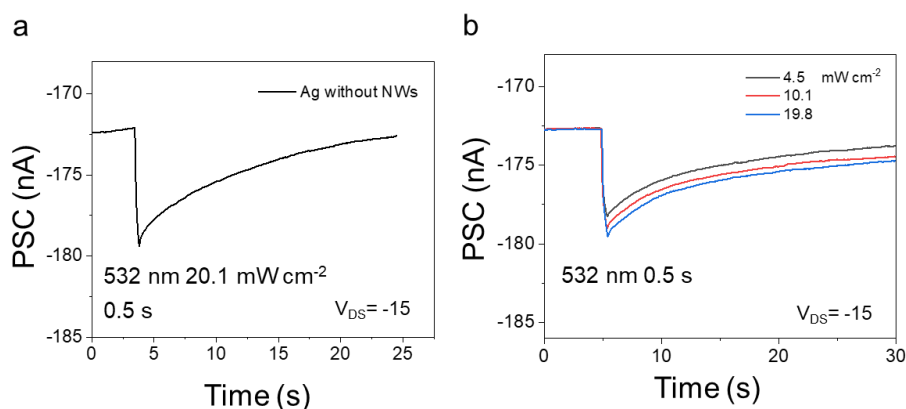
Magnified optical image of each corresponding unit. **e**, Transfer characteristics of the 25 randomly selected devices.
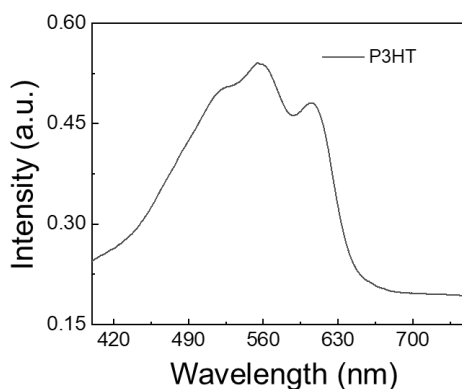

Supplementary Fig. 15| Absorption spectrum of the self-assembled P3HT film.
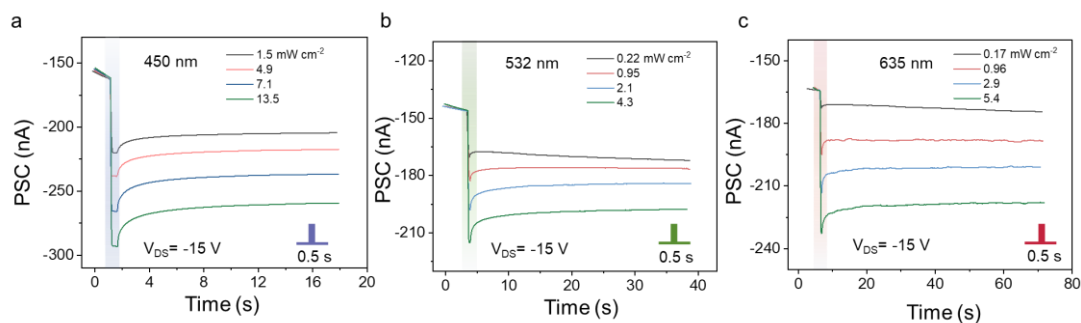

Supplementary Fig. 16| Typical synaptic plasticity of EPSC for the optoelectronic synapse with **a**, 450 nm **b**, 532 nm, and **c**, 635 nm single pulses (0.5 s pulse width).


Supplementary Fig. 17| **a**, EPSC of the self-assembled P3HT film with Ag source/drain electrodes. **b**, EPSC of the self-assembled P3HT film with Ag source/drain electrodes and different light powers.
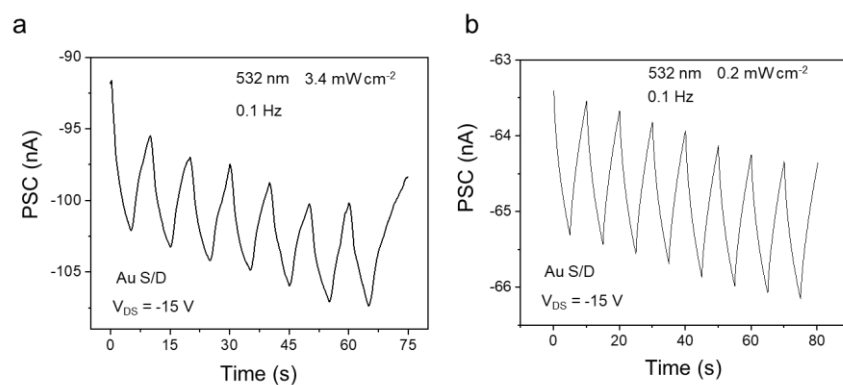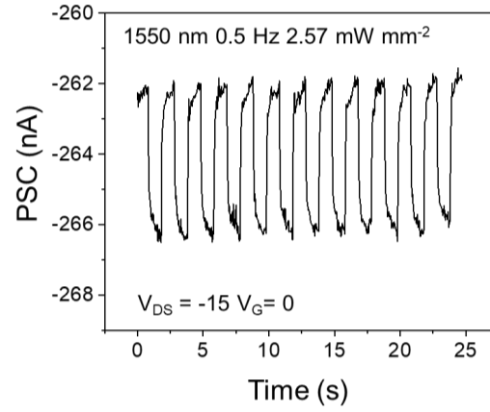
Supplementary Fig. 18| Photoresponse of the self-assembled P3HT film with Au electrodes under different power irradiations. **a**, 3.4 mW cm^{-2}. **b**, 0.2 mW cm^{-2}.
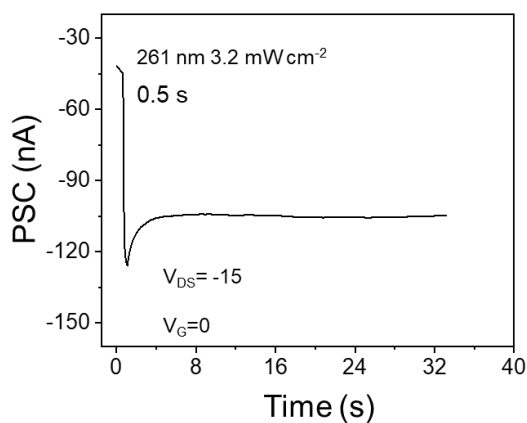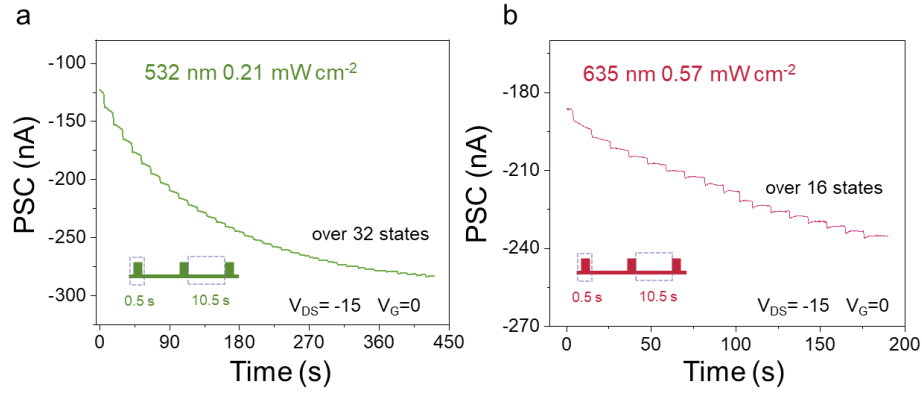

**Supplementary Fig. 19|** Non-volatile optical signal storage with 261 nm single pulse.
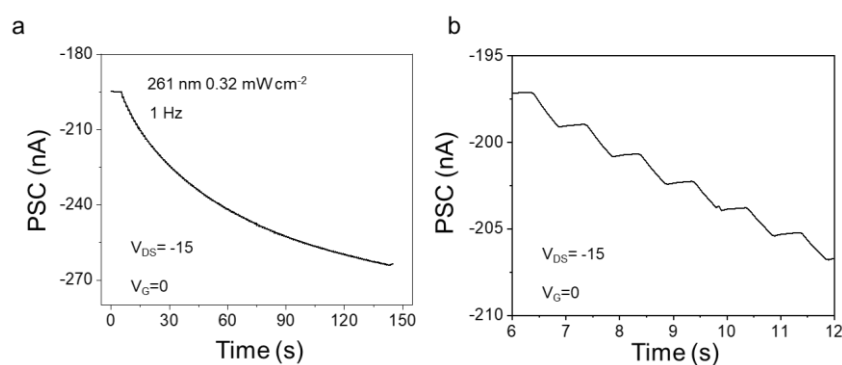
Supplementary Fig. 20| a, Linear changes in postsynaptic synaptic weights and surpassing 128 memory states under continuous 261 nm pulses (1 Hz). **b**, Enlarged view of the corresponding details.

**Supplementary Fig. 21|** Photoresponse of the synaptic device in the infrared range.

**Supplementary Fig. 22|** In-memory sensing and computing process spikes with a, 532 nm and b, 632 nm.
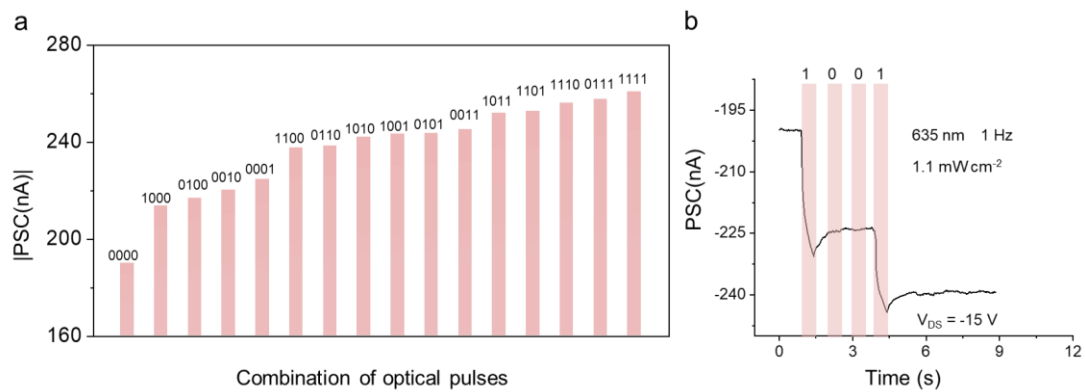
Supplementary Fig. 23| a, Responses to 4-bit light stimulation. **b**, Encoding process for '1001' stimulation. The current is extracted from the value of the current at the end of the last four 1 Hz cycles.
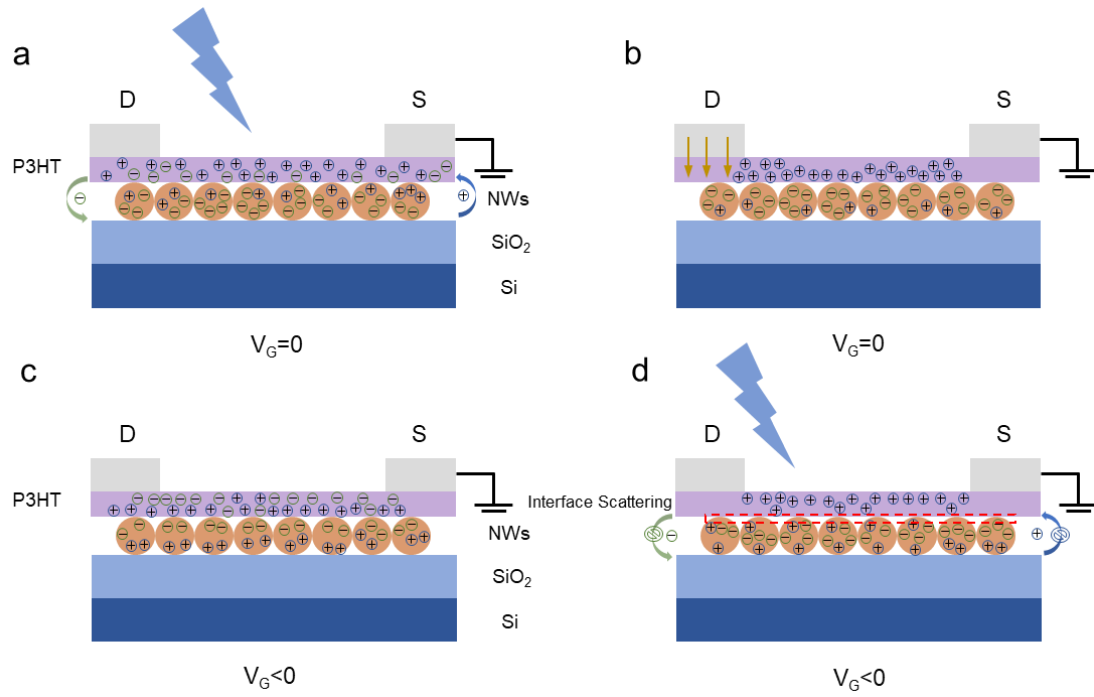
Supplementary Fig. 24| Schematic diagram of the device operation principle at different V_G. **a**, $V_G = 0$ and light is applied. Photogenerated carriers are separated and transported in the presence of the heterojunctions. **b**, $V_G = 0$ and the light stops. Strong carrier retention effect under the action of the dual heterojunctions. **c**, $V_G < 0$ and no light. Charge induction within the semiconductors and carrier aggregation in the P-type channel at negative gate voltage. **d**, $V_G < 0$ and light is applied. Carriers recombination and interfacial scattering within the semiconductors cause negative photoconductivity effect.
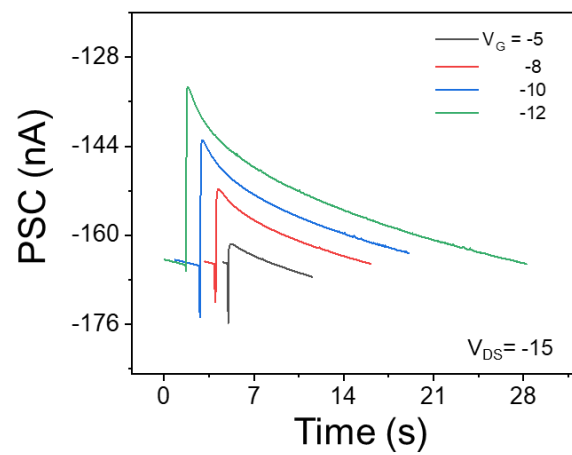
Supplementary Fig. 25| EPSC of the synaptic devices induced by the different V_G spikes.
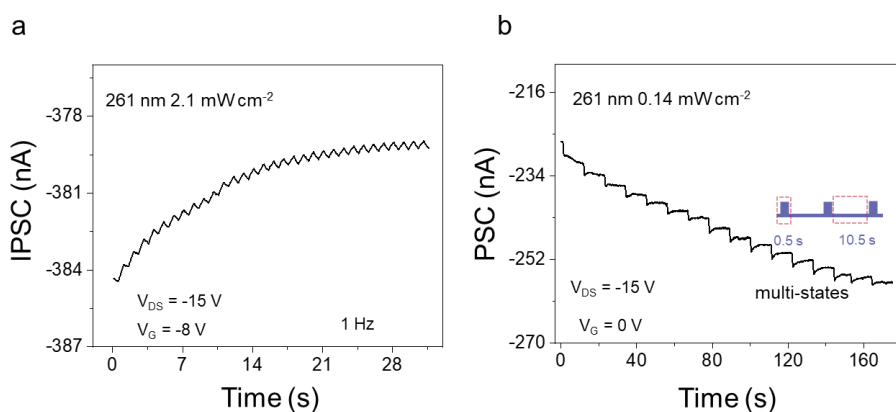
Supplementary Fig. 26| a, NPC, and **b**, PPC long-term plasticity realized by 261 nm pulses and different V_G.
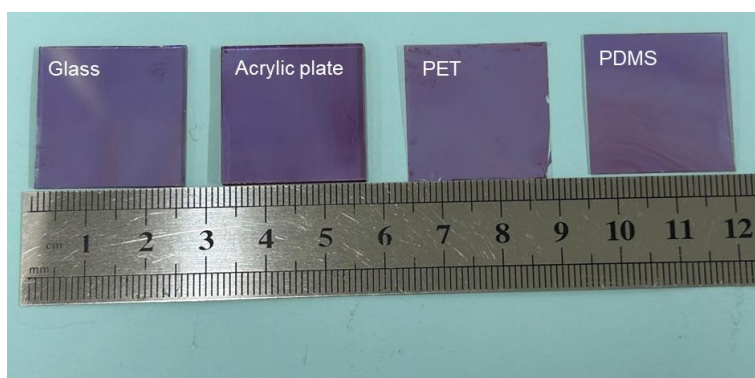
Supplementary Fig. 27| The self-assembled P3HT films were physically transferred to various 2.5×2.5 cm substrates, including glass, acrylic plate, PET, and PDMS.
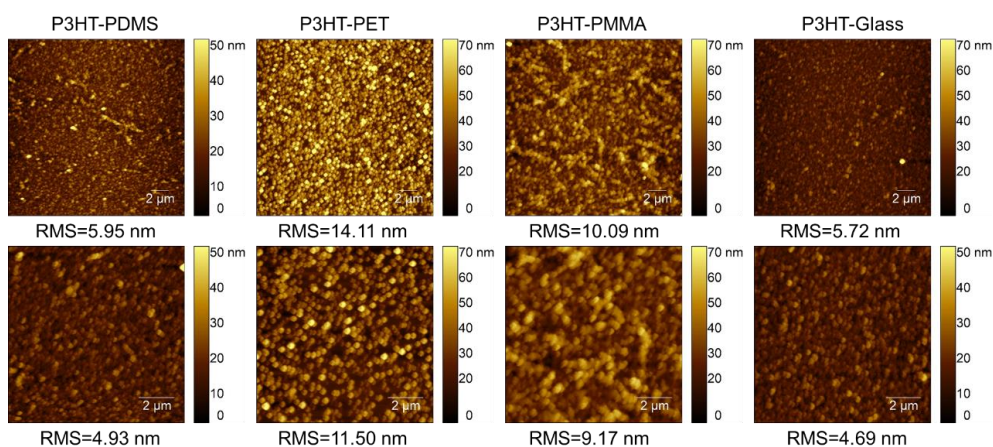
Supplementary Fig. 28| AFM images of the self-assembled P3HT films on different substrates.
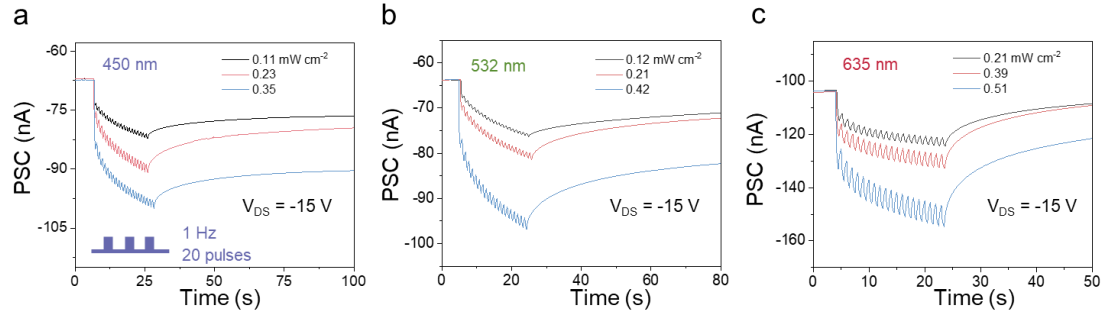
Supplementary Fig. 29| Long-term plasticity of the synaptic devices with **a**, 450 nm, **b**, 532 nm, and **c**, 635 nm 20 consecutive presynaptic light spikes (1 Hz) on the PI substrate.
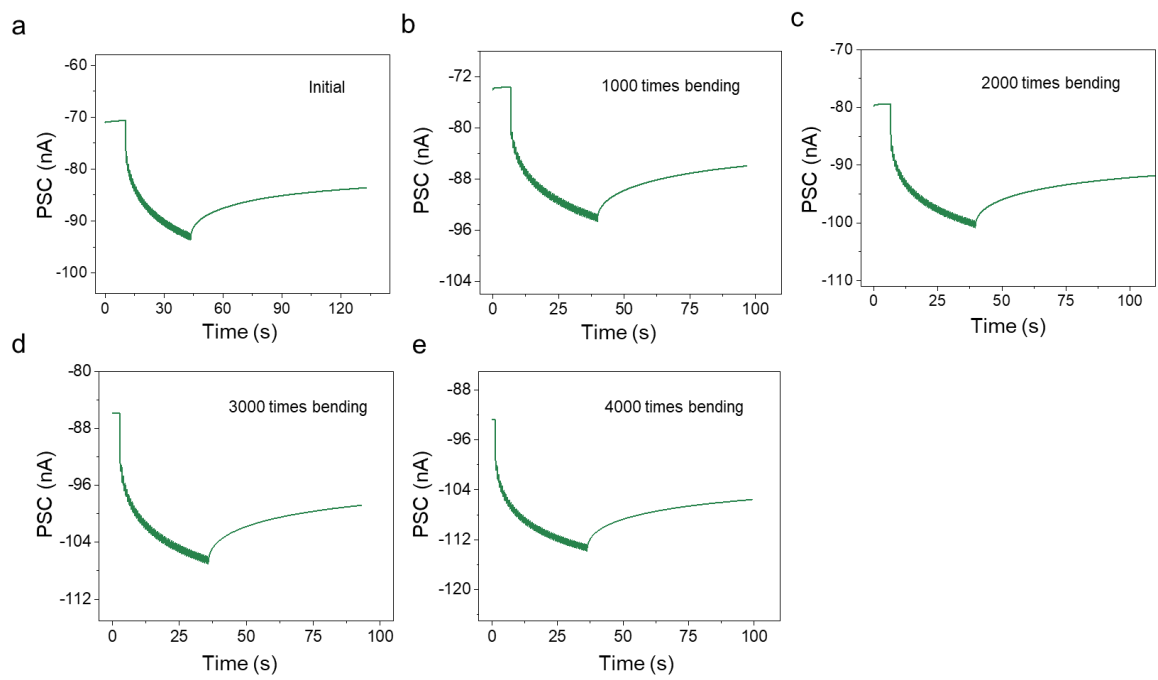
Supplementary Fig. 30| Bending test of the synaptic devices on a PI substrate with an **a**, initial, **b**, 1000 times, **c,** 2000 times, **d**, 3000 times, and **e**, 4000 times bending.
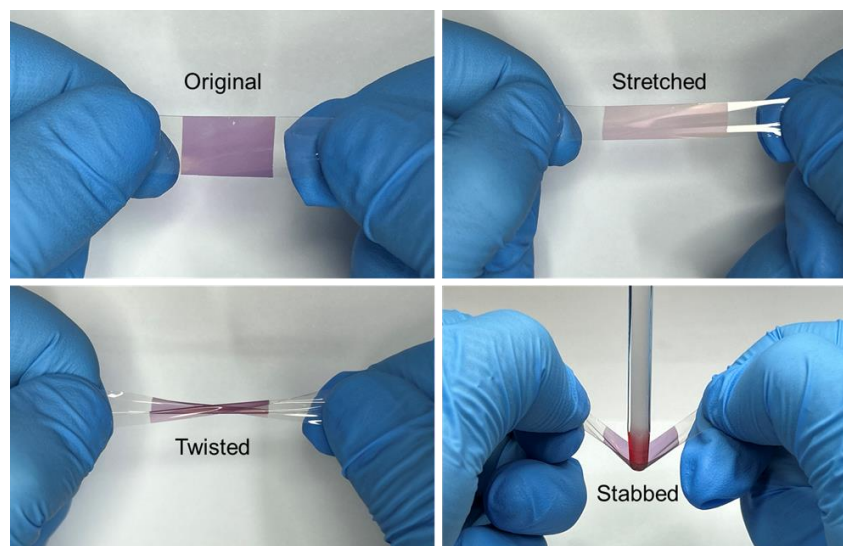
Supplementary Fig. 31| Self-assembled P3HT films on the PDMS substrate with different deformations such as stretching, twisting, and poking.
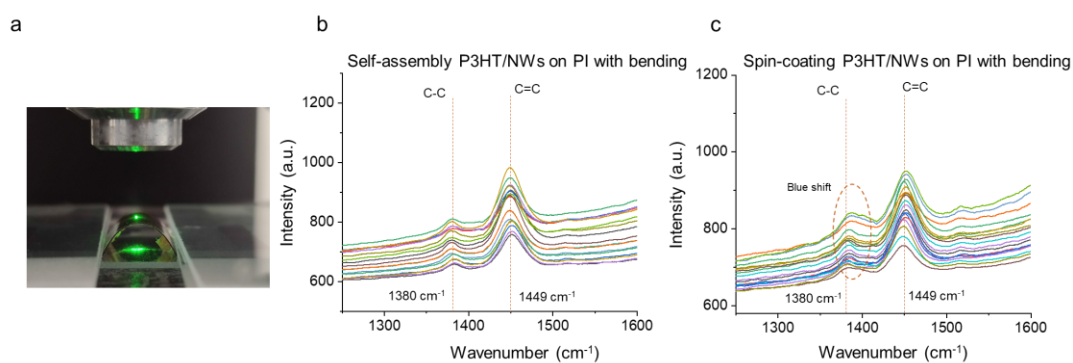
Supplementary Fig. 32| a, Photograph of the self-assembled P3HT/NWs film on the bending PI substrate. The Raman peaks of **b**, the self-assembly P3HT/NWs film on the bending PI substrate, and **c**, the spin-coated P3HT/NWs film on the bending PI substrate.
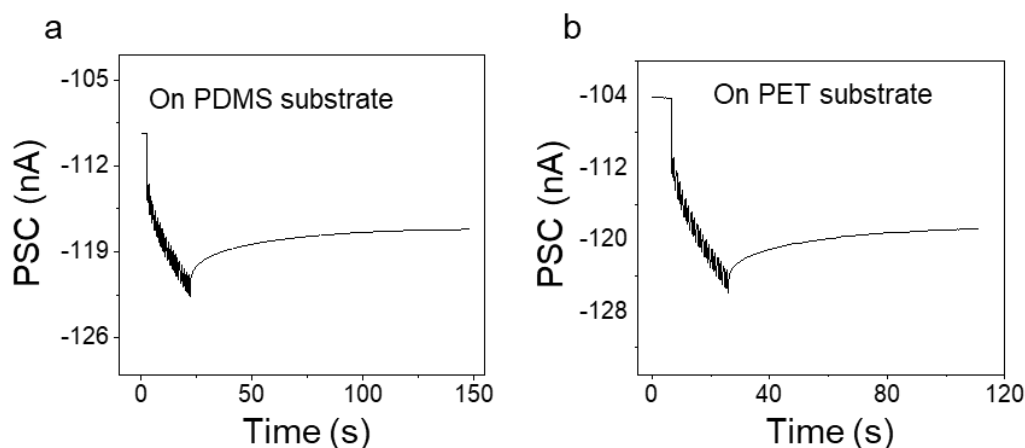
Supplementary Fig. 33| The performance of the flexible device on different substrates. **a**, PDMS, and **b**, PET. (532 nm, 1.5 Hz, 0.25 mW cm^{-2}).
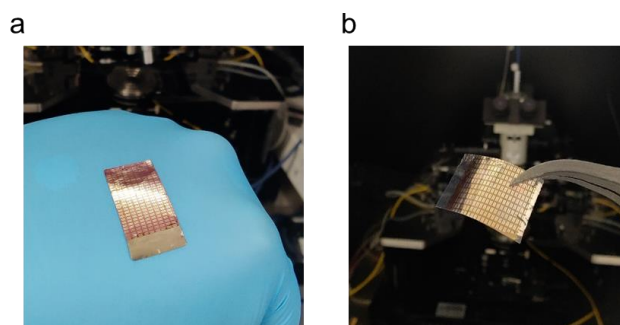
Supplementary Fig. 34| Wearable synaptic devices arrays on the PET substrate. **a**, Paste on glove. **b**, The array can be picked up and integrated onto other substrates.
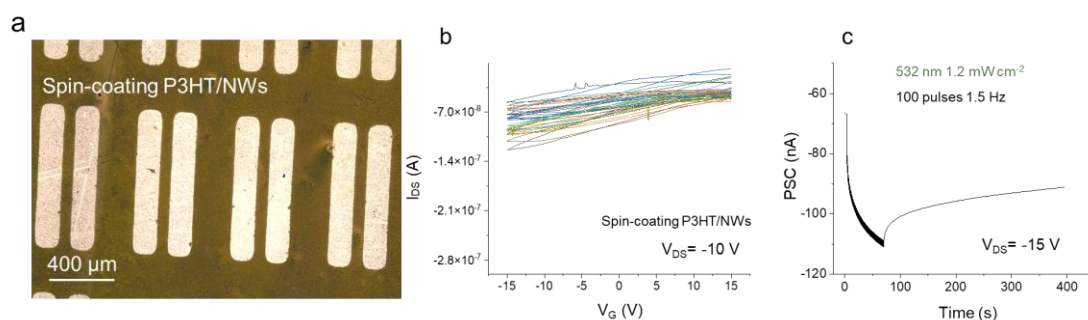
Supplementary Fig. 35| a, Photograph of the spin-coated P3HT/NWs devices on the Si/SiO$_2$ substrate (Ag S/D electrodes). **b**, Transfer characteristic curves of the randomly selected 18 devices with the spin-coated P3HT film. **c**, EPSC of the spin-coated P3HT film.
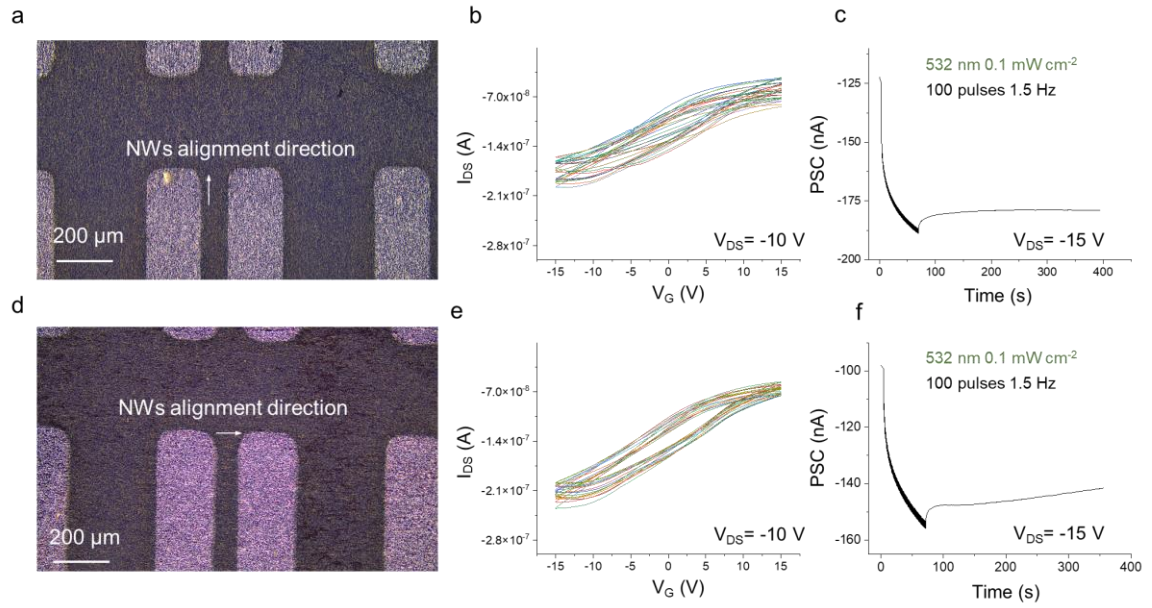
Supplementary Fig. 36| a, **d**, Optical microscope images (dark field) of the self-assembled P3HT/NW devices, different NW alignment directions, on the Si/SiO$_2$ substrate with Ag S/D electrodes. **b**, **e**, Transfer characteristics of the 17 randomly selected devices. **c**, **f**, Corresponding EPSC with the 532 nm light pulses.
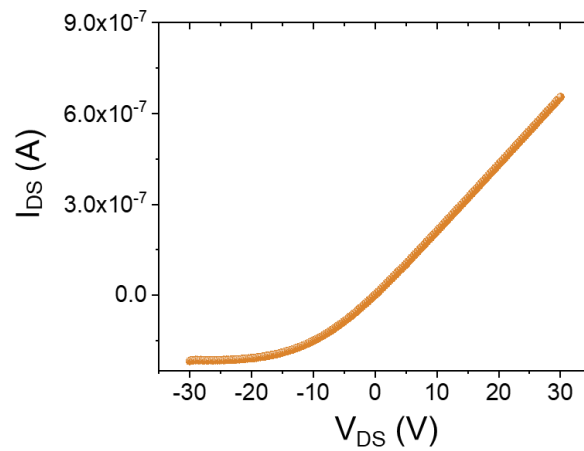
Supplementary Fig. 37| I-V curves of the synaptic devices demonstrate the existence of a Schottky barrier.
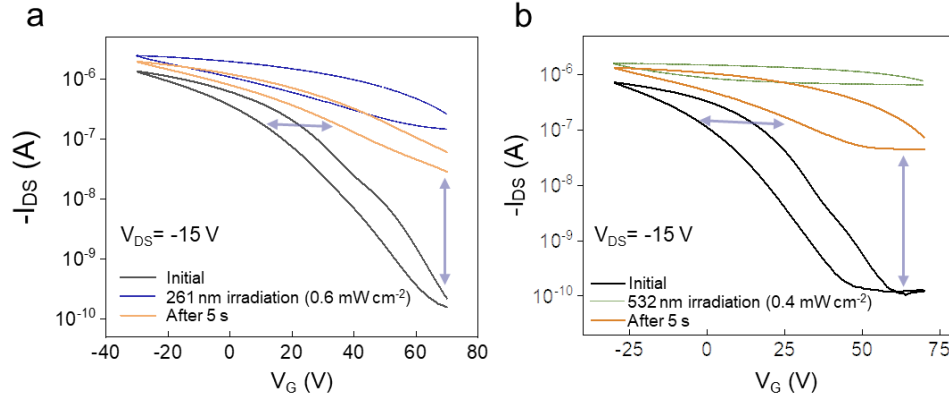
Supplementary Fig. 38| Transfer characteristic curves of the three-terminal synaptic devices before and after light irradiation with **a**, 261 nm and **b**, 532 nm light.
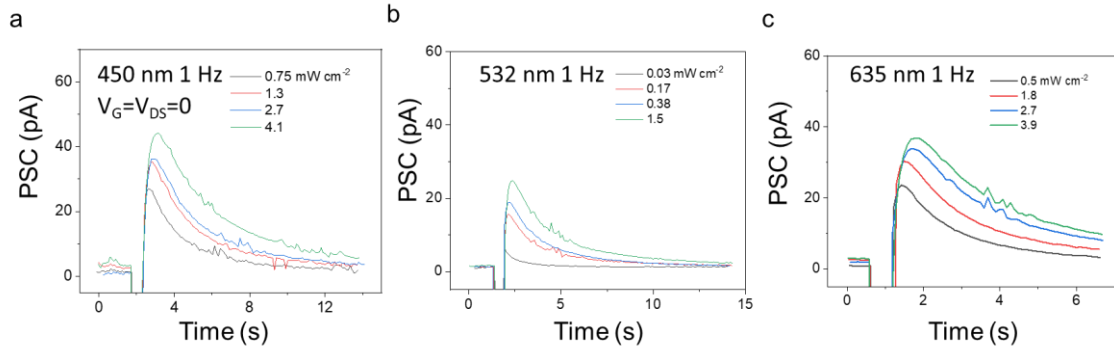
Supplementary Fig. 39| Different synaptic behaviors, including EPSC and SIDP, were successfully mimicked with a, 450, b, 532, and **c**, 635 nm single light pulse under V_G and $V_{DS} = 0$.
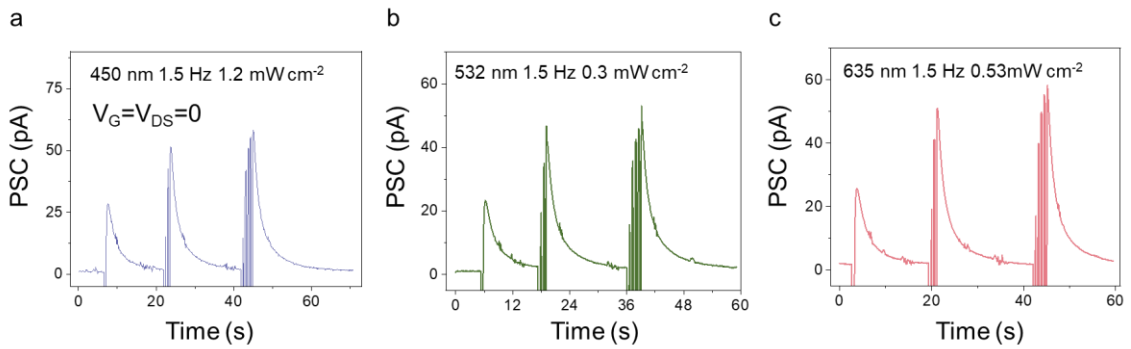
Supplementary Fig. 40| SNDP of synaptic devices with **a**, 450, **b**, 532, and **c**, 635 nm (1, 3, 5 light pulses) under V_G and $V_{DS} = 0$.
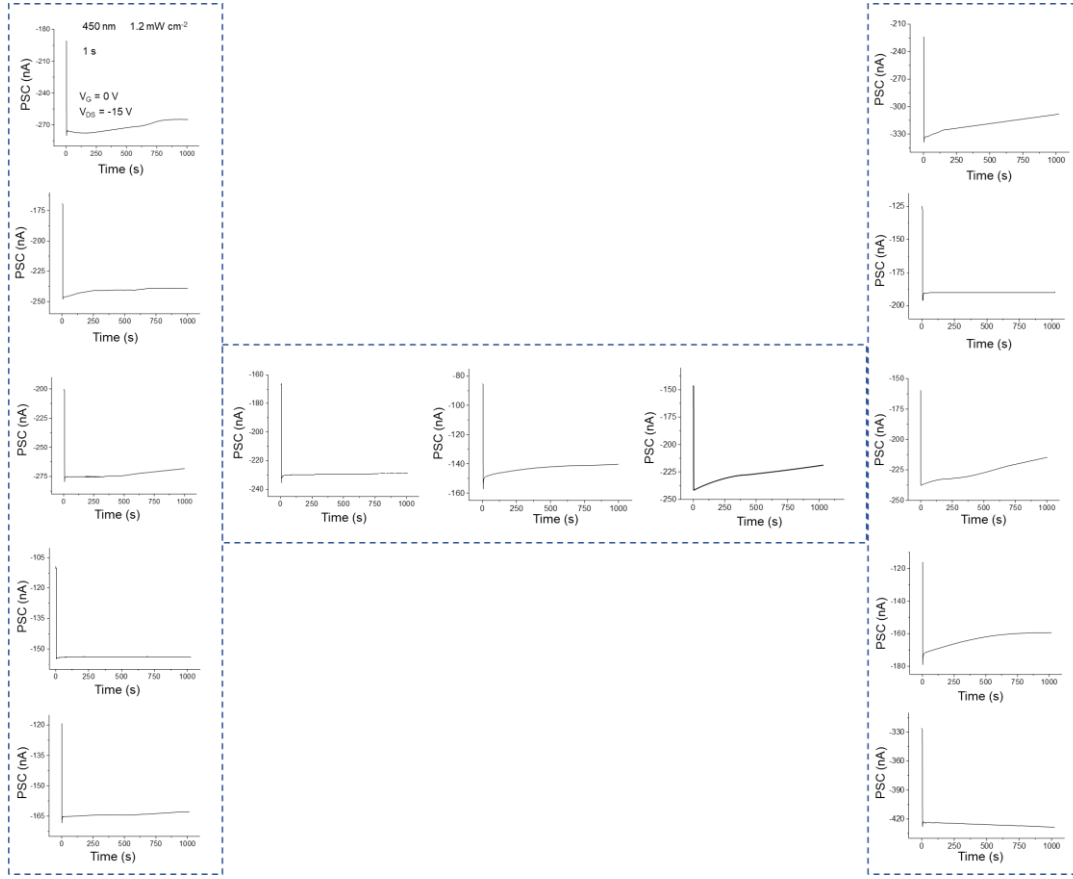
Supplementary Fig. 41| Non-volatile optical information storage of the device after a single light pulse irradiation. The H-shaped feature contains 13 individually tested devices, while the rest of the devices gave stable dark currents.
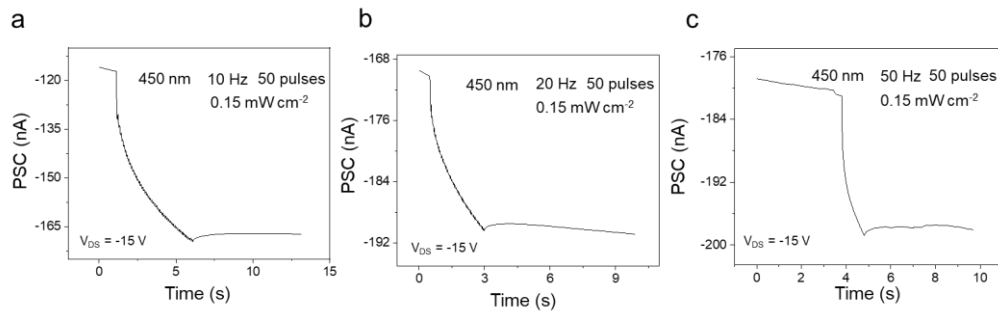
Supplementary Fig. 42| EPSC with **a**, 10 Hz, **b**, 20 Hz, and **c**, 50 Hz 450 nm light pulses.

Supplementary Fig. 43| Schematic of the training and testing processes for color recognition in reservoir computing. Inspired by this, we convert the different colors into three bands of color light pulses. During the training process, we used a self-built dataset containing 104 color data and 6 types of labels. For input to the neural network, we use a $3 \times 200 \times 6$ neural network and use the Relu function as the activation function of the hidden layer, and the output labels represent the 6 colors. When recognizing a moving object's color, we input the color data of the moving subject into the device reservoir, sample the device's state to operate directly with the trained weight network, and obtain the target label value. (Blue process arrows represent the training process, yellow process arrows represent the testing process, and blue-yellow process arrows represent the process shared in training and testing.)
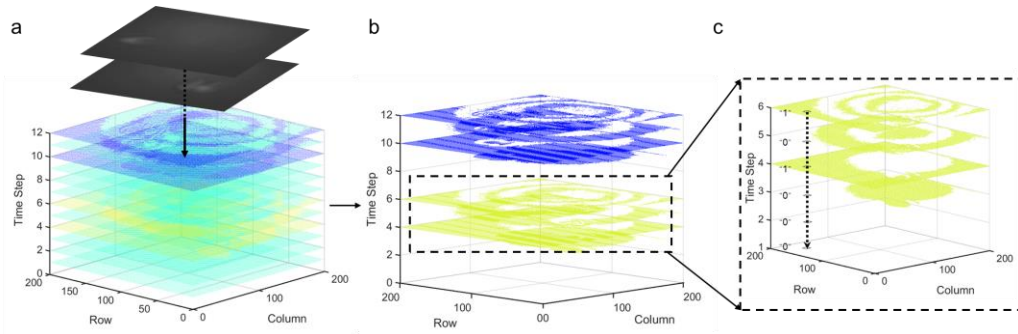

Supplementary Fig. 44| a, The sample input of the selected adjacent frame stream is displayed in a spatiotemporal view. **b**, The efficient spatiotemporal view of input extracted without "0". **c**, The details of input for positive photoconductance mode.
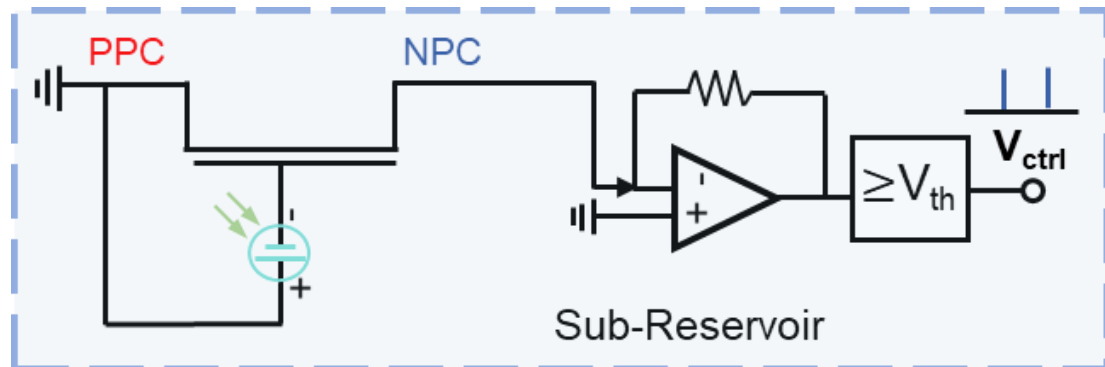
Supplementary Fig. 45| Schematic of the sub-reservoir circuit for motion detection. NPC and PPC in the blue boxed circuit diagram correspond to G_+ and G_- in Fig. 6b in

the manuscript, respectively. At the end of every two frames of image input, the voltage state of each pixel in the array is obtained. During this process, the first frame is fed into PPC, causing a positive change in the device's voltage state. The second frame is provided in NPC. The device's voltage state inherits the state of the previous frame and changes negatively. After two frames, the device's voltage state is acquired. Input the third frame into NPC, the fourth frame into PPC, and so on. When the input light pulse of the previous and subsequent frames remains unchanged, the acquired voltage state approaches 0. Adjusting the threshold (V_{th}) in the threshold setting circuit of the sub-reservoir circuit makes it possible to avoid the interference of environmental noise. Moreover, by modifying different thresholds, the outcome of motion recognition varies. When the input light pulses differ between the former and the latter, the acquired voltage state is significantly higher than the 0 potential. The current is only output when it exceeds the threshold value, realizing the recognition of moving objects.
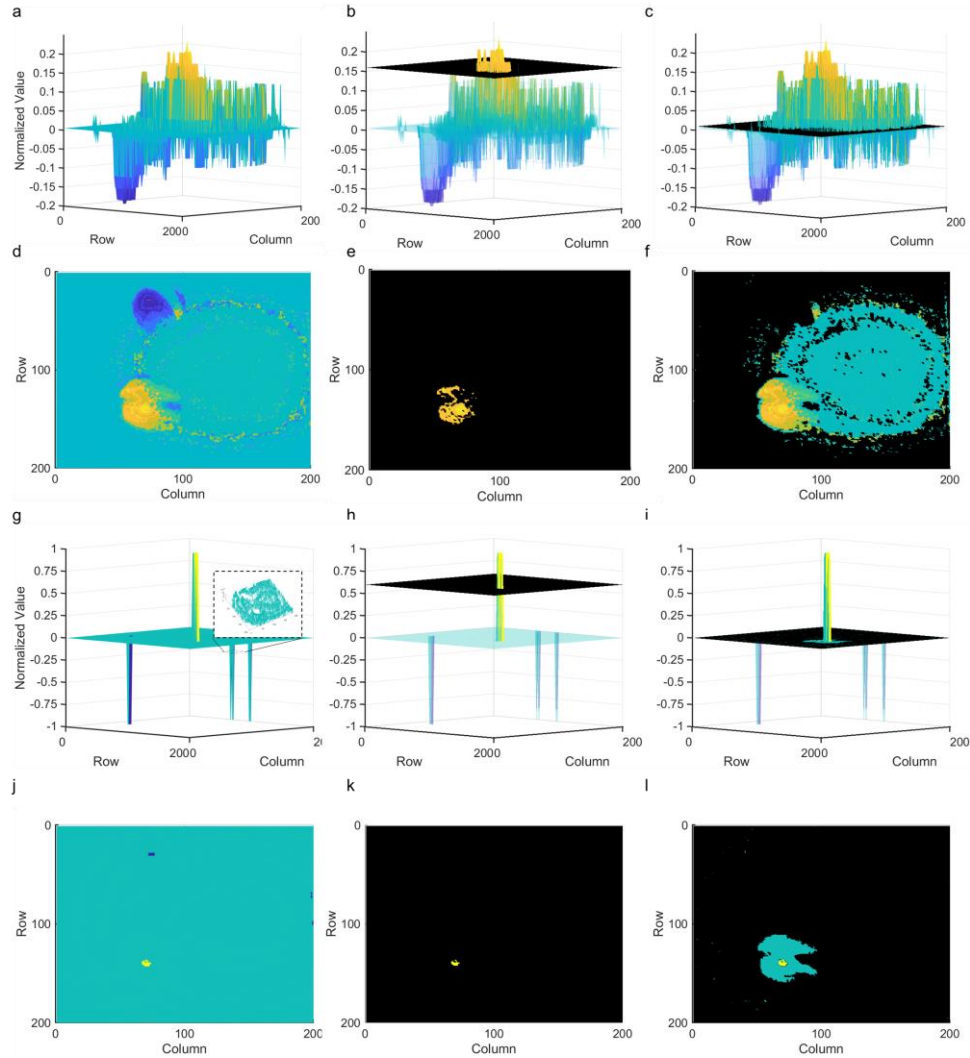
Supplementary Fig. 46| The normalized final state in 3D view generated by our method **a**, without threshold plane, and **b**, **c**, with different threshold planes. **d-f**, The

2D result corresponding to **a-c**. The normalized final state in 3D view generated by conventional method **g**, without 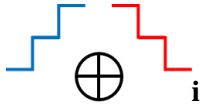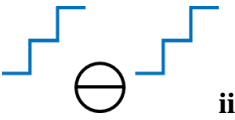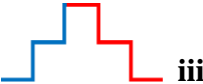threshold plane (with a zoomed-in detail plane in it) and **h**, **i**, with different threshold planes. **j-l**, The 2D result corresponding to **g-i**.

As a comparison, we utilize the conventional method of current subtraction to conduct motion detection and observe the effect (as a demonstration, we uniformly normalize the data to the [-1,1] interval for better visualization of the effect). We still select the two frames in Supplementary Fig. 43. Subsequently, **a** and **g** represent the current state after inputting two frames. **a** illustrates the proposed method, and **g** depicts the conventional method for comparison. In the case of very low ambient brightness, the object's brightness merged with the surroundings' brightness. The traditional method does not work well (shown in **j**). By contrast, the method proposed in this paper demonstrates the result of motion detection (shown in **d**). Based on this, the detection effectiveness is enhanced by applying thresholds, **b, c** and **h, i** are applied with different thresholds, and it can be observed that there is a noticeable change in the results of motion detection.
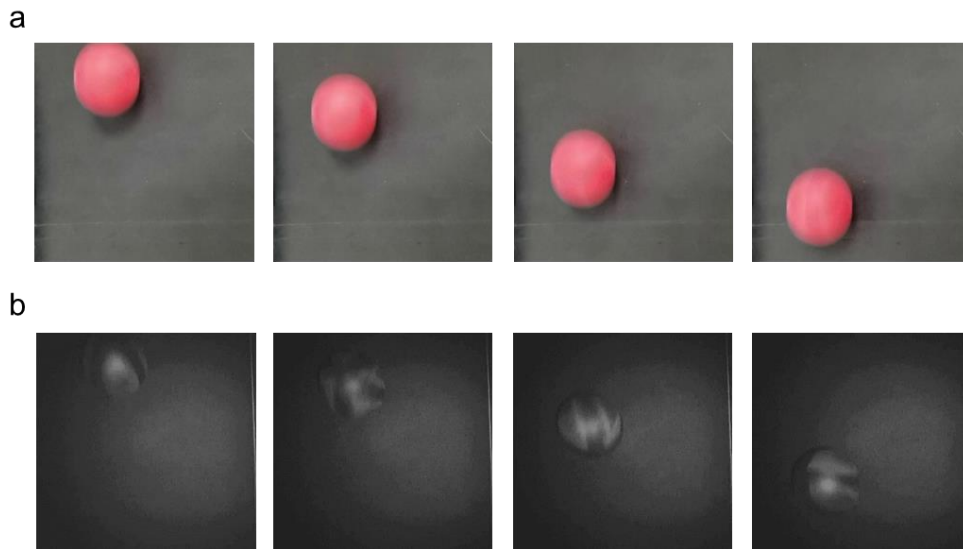
Supplementary Fig. 47| Screenshot of the videos filmed by the **a**, normal camera (visible range) and **b**, UV camera, respectively.

Supplementary Table 1| The comparisons with related works in the treatment of positive and negative photoconductance.

Works	Channel	Method	$\Delta N = \Delta P$	Additional circuit	Device array parameter
1[1]	P/N junction	i	False	Step function	2n
2[2]	P/N junction	As above	False	Step function	2n
3[3]	P/N junction	As above	False	Step function	2n
4[4]	P channel	ii	/	Step function	2n
Our work	P/N junction	iii	False	Threshold circuit	n

Notes:

i Adding currents produced by positive and negative photoconductive device arrays in neighboring frames.

ii Subtract of currents produced by arrays of 2 positive photoconductive devices respectively in neighboring frames.

iii Direct currents produced by the dynamic characteristics of arrays of devices with positive and negative photoconductance in neighboring frames.

For traditional motion detection, such as literature 1 to 3 listed in Supplementary Table 1, the principle of current addition is first applied to the positive and negative photoconductive device array outputs. In the case of motion detection, as the fourth reference discussed in Supplementary Table 1, where there is no negative

photoconductive response of the device, differential amplifiers are used to achieve motion detection by subtracting the currents from the two positive photoconductive device array outputs. The methodology used in our work is based on the dynamic characteristic of a single device, utilizing the non-volatile and positive and negative photoconductors of the device to perform current addition and subtraction for motion detection directly. Then, the effect of motion detection can be altered by employing the threshold regulation method. Moreover, for the first two methods, an equal magnitude of weight change of NPC and PPC can be required. However, even if they are not the same, it can be achieved by outputting to a Step Function. Based on the former, we have improved the method in our work. With a threshold circuit, it can be effective for motion detection in the case of a fixed difference value. In the same situation, the number of devices in the array in this method is half that of the previous two methods. This reduction is because the previous two methods usually require the currents generated by two neighboring frames to be processed for addition or subtraction.

Supplementary references:

1   Zhang, Z. et al. All-in-one two-dimensional retinomorphic hardware device for motion detection and recognition. *Nat. Nanotechnol.* **17**, 27-32 (2022).
2   Zhu, X. et al. High-contrast bidirectional optoelectronic synapses based on 2D molecular crystal heterojunctions for motion detection. *Adv. Mater.* **35**, 2301468 (2023).
3   Pang, X. et al. Non-volatile rippled-assisted optoelectronic array for all-day motion detection and recognition. *Nat. Commun.* **15**, 1613 (2024).
4   Jiang, T. et al. Tetrachromatic vision-inspired neuromorphic sensors with ultraweak ultraviolet detection. *Nat. Commun.* **14**, 2281 (2023).