Electronic supporting Information

Non-Linear Two-Dimensional Float Gate Transistor as Lateral Inhibitory Synapse for Retinal Early Visual Processing

Man Hu,‡[a] Jun Yu,‡[a] Yangyang Chen,[b] Siqi Wang,[b] Boyi Dong,[b] Han Wang,[a] Yuhui He,[b] Ying Ma,[a] Fuwei Zhuge,*[a] Tianyou Zhai*[a]

[a.] State Key Laboratory of Material Processing and Die and Mould Technology, School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, P. R. China

[b.] School of Optoelectronic and Information, Huazhong University of Science and Technology, Wuhan 430074, P. R. China

*Corresponding Authors. E-mail: zhugefw@hust.edu.cn and zhaity@hust.edu.cn.

‡ These authors contributed equally to this work.

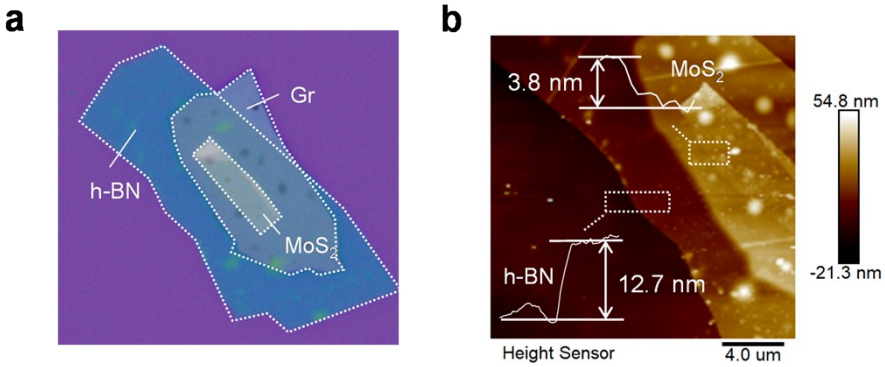

Fig. S1. Thickness characterization of Gr/h-BN/MoS_2 heterostructure: (a) optical image of the 2D vdW stacked heterostructures of MoS_2/h-BN/Gr on SiO_2/Si substrate. (b) thickness of the h-BN and MoS_2 flakes are determined by atomic force microscopy to be 12.7 and 3.8 nm respectively.

Note 1: Operation characteristic of subthreshold transistor and float gate transistor

At such subthreshold operation, the drain to source current in transistor becomes non-linear depending on the gate to channel coupling, thus displays rectifying behavior. The I_{DS} current in a subthreshold transistor is described by:

$$I_{DS} = I_0 exp\left(\frac{eV_{GS}}{nk_BT}\right)\left[1 - exp\left(-\frac{e\left(V_{DS} - I_{DS}R_s\right)}{nk_BT}\right)\right] \qquad (1)$$

$$I_0 = AA^* T^2 exp\left(\frac{-e\phi_B}{k_BT}\right) \qquad (2)$$

where n is the ideal factor, R_s is the serial resistance, I_0 is the thermionic emission current at the Schottky contact with the barrier ϕ_B, A is contact area, A^* is the Richardson constant.

Thus, the reverse saturation current not only depends on the Schottky barrier ϕ_B at electrode contact, but also influenced by gate to source coupling. Since the contact barrier for MoS_2/Cr/Au is low, usual MoS_2 transistor does not manifest apparent rectifying behavior. Here, in our work, the function of ZnPc is essential, which modifies the threshold of MoS_2 channel by depleting its electron concentration. This leads to remarkable decrease of contact potential difference (CPD) and extracted Schottky barrier displayed in Fig. 1e and Fig. 2f.

In the case of float gate transistor, the float gate potential V_{FG} is influenced by the charge trapping and applied voltage bias at channel. Because of the large lateral resistance of the channel in subthreshold regime, in Fig. S2, we considered a simplified case by including only drain/source electrode coupling to float gate, and float gate to control gate. In such case, we could write:

$$V_{FG} = \frac{Q_{trap} + V_{DS}C_{FG-D}}{C_{FG-S} + C_{FG-CG} + C_{FG-D}} = V_{FG,0} + V_{DS}\frac{1}{\frac{C_{FG-S} + C_{FG-CG}}{C_{FG-D}} + 1} \qquad (3)$$

In our case, because of the extended float gate configuration, $C_{FG-S} + C_{FG-CG} \gg C_{FG-D}$, and the float gate to control gate coupling ratio approaches 1. In this case, as long as charge tunneling does not occur when sweeping V_{DS}, the float gate potential only slightly increases with V_{DS}. Thus, the rectification behavior based on drain to float gate couple remains.

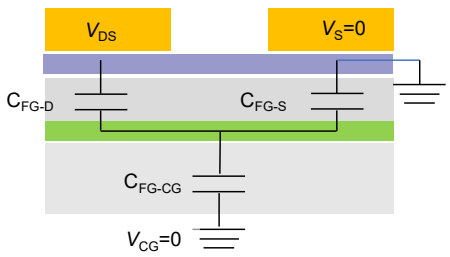

Fig. S2. Simplified equivalent circuit for the capacitance coupling in the float gate transistor.

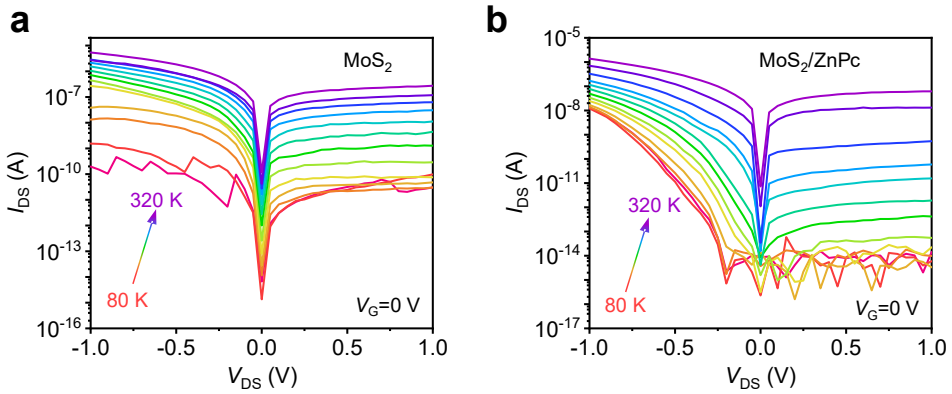

Fig. S3. Temperature dependent measurement of initial MoS_2 transistor and the one after 40 min's ZnPc molecule processing: (a), (b) I_{DS}~V_{DS} curves at different temperature from 80 K to 320 K at V_G=0 V, both the reverse saturation current and forward current increases with temperature, as described by equation 1.

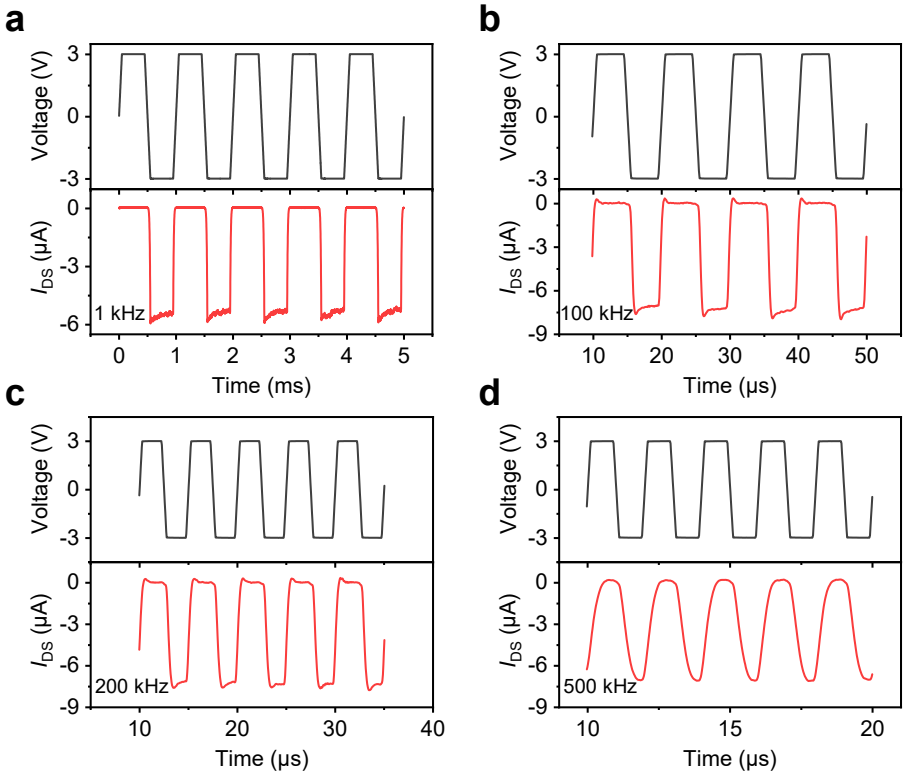

Fig. S4. Frequency dependent rectification to square waveform input at drain: (a), (b), (c), (d) 1, 100, 200 and 500 kHz. The graphene electrode was set as ground for a fixed gate to source coupling.

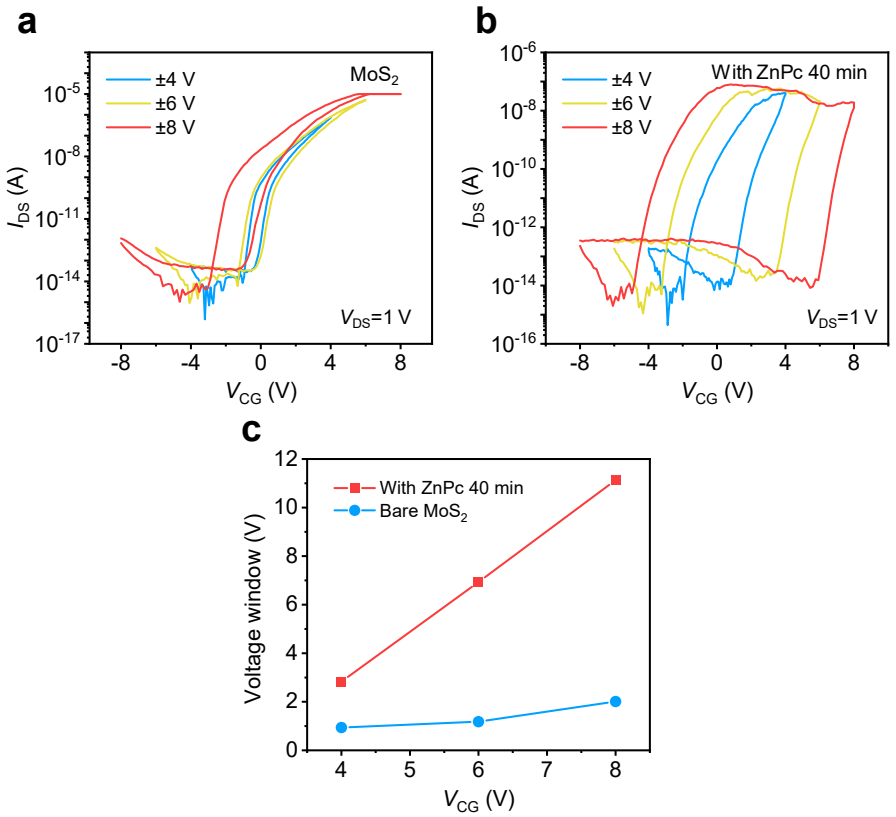

Fig. S5. Memory window of initial MoS_2 FGFET and the one with ZnPc decoration: (a), (b) hysteresis loop of device under different sweep range of V_{CG}. (c) comparison of the memory window extracted from (a), (b). It is apparent that introducing ZnPc molecules significantly improves the memory window. This is interpreted as a result of lowered hole tunneling barrier by depleting the MoS_2 channel.

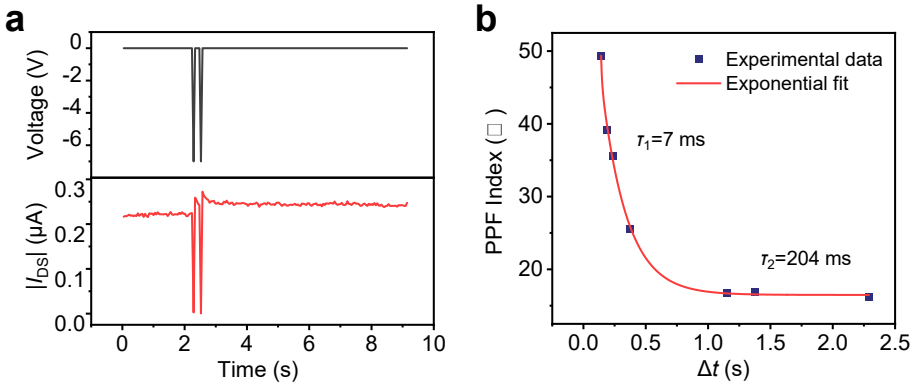

Fig. S6. Paired pulse facilitation (PPF) characteristic of the device: (a) PPF behavior of the present FGFET based on short-term relaxation behavior. (b) pre-synapse spikes (V_{CG}=-7 V, t_w=50 ms) paired with different intervals were applied from the control gate. The extracted PPF index exhibited apparent interval dependent behavior, which follows a classic exponential law as those observed in biology synapses: $y = C_1 exp(-t/\tau_1) + C_2 \exp(-t/\tau_2)$. The time constants of τ_1 and τ_2 are fitted to be 7 ms and 204 ms, respectively.

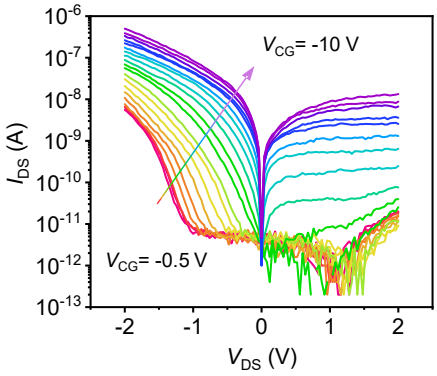

Fig. S7. The logarithmic plot of current-voltage characteristic of FGFET from V_{DS}=-2 V to 2 V.

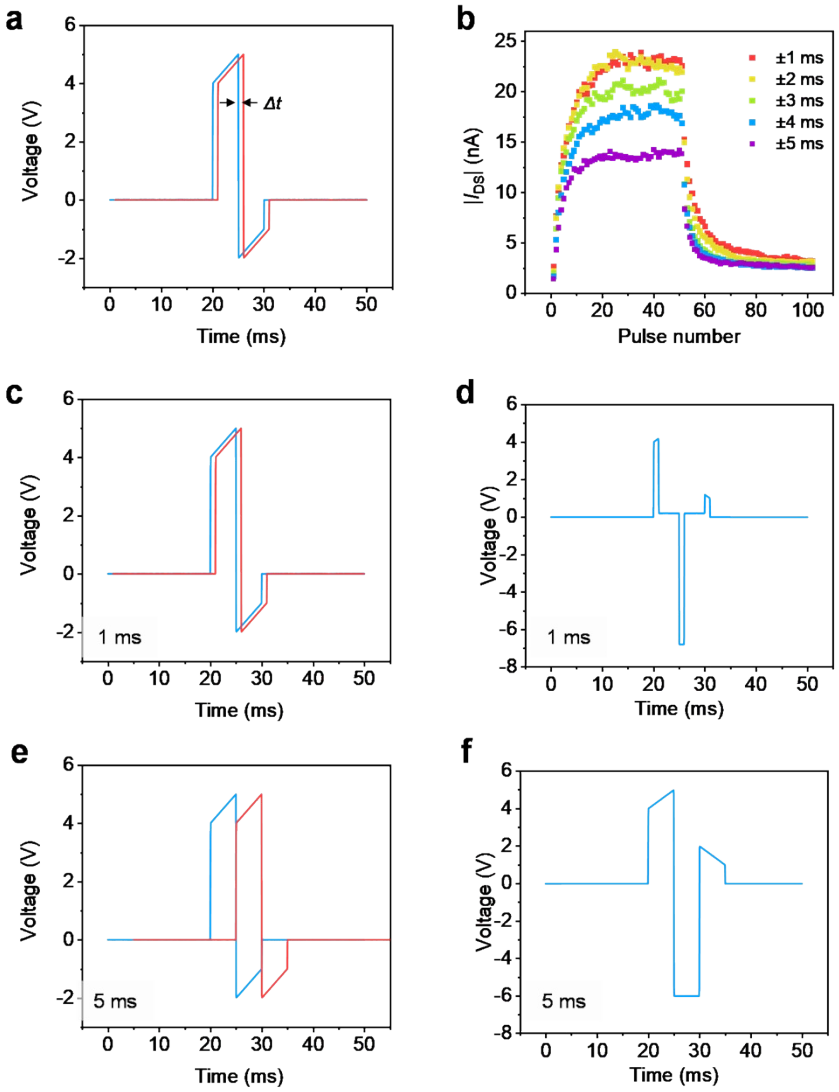

Fig. S8. Long term potentiation and depression of synapse conductance via spike time dependent plasticity: (a) the waveform of spike pairs applied in experiment. (b) the observed read current change upon different pulse number of spike pairs. The synapse weight is continuously tuned in this way depending on the spike time sequency between pre- and post-neurons. (c), (d) the effective waveform of spike pairs and equivalent spike with 1 ms and (e), (f) 5ms interval, the time interval influences the peak voltage drop and the maximum of attained conductance upon repeated spike pairs.

Note 2: Numerical simulation of the effective convolution kernel

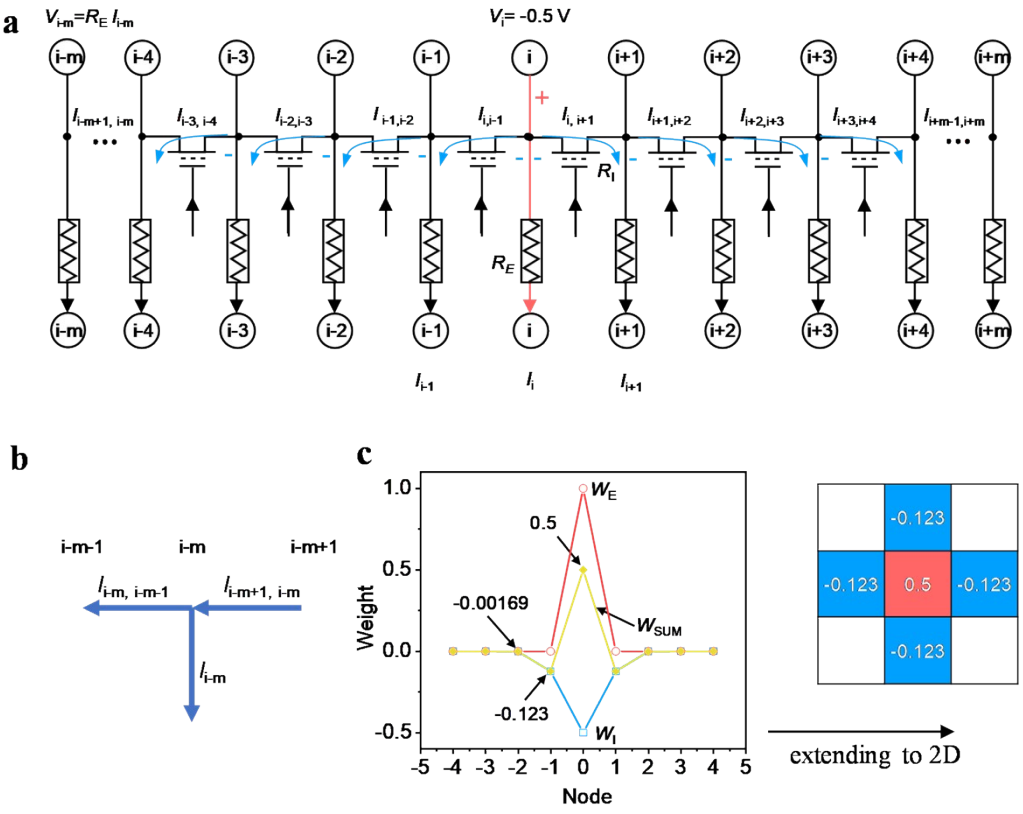

Fig. S9. (a) illustration of a one-dimensional network consisted of both feedforward and lateral signal transmission. (b) the current direction at the node i-m. (c) simulated weight distribution using the current-voltage characteristic of transistor in Fig. 5c and R_E=1.1 GΩ, V_P=1 V, V_N=-0.5 V, when extending in a two-dimensional network, the function of the network could be approximated by applying a 3×3 kernel.

Fig. S9a illustrate a simplified one-dimensional network configuration, with voltage pulse applied at a certain pixel with marked address i. Feedforward connections are made between input and output layer, while lateral connections are introduced for nearest neighbor pixels. Due to the negative-pass rectification behavior of the lateral transistor, the positive part of input waveform only transmits forwardly, and the positive (excitatory) weight could be simply calculated according to:

$$I_i = \frac{V_i}{R_E} \qquad (4)$$

However, for negative part, both feed-forward and lateral transmission co-exist, and the current flow to output neurons depends on the node's distance to input i. Due to the non-linear conductance of lateral transistors, the voltage distribution in network for negative part could not be analytically solved. To drive the negative (inhibitory) weight distribution, we adopted numerical calculation methods to estimate the current transmission at output layer, using the current-voltage characteristic of the transistor displayed in Fig. 5c. In the example shown in Fig. 5, parameters values were taken as R_E=1.1 GΩ, V_P=1 V, V_N=-0.5 V.

For simplicity, the current-voltage characteristic of the transistor was approximated in equation 5 by ignoring serial resistance R_s:

$$I_{DS} = -I_0 exp\left(\frac{eV_{GS}}{nk_BT}\right)\left[1 - exp\left(-\frac{eV_{DS}}{nk_BT}\right)\right] \qquad (5)$$

For the considered input waveform with negative magnitude less than 0.5V, this still yield accurate fitting to the current-voltage characteristic using $I_0 = -1.49 \times 10^{-14}$ A, n=2.06.

Assuming a 1x9 network with voltage pulse applied at the center node i=5. The voltage distribution in the network was calculated according to:

$$I_{i-m+1,\ i-m} = I_{i-m,\ i-m-1} + I_{i-m} \qquad (6)$$

$$I_{i-m,\ i-m-1} = f\left(V_{i-m-1}, V_{i-m}\right) \qquad (7)$$

$$I_{i-m} = \frac{V_{i-m}}{R_E} \qquad (8)$$

where $I_{i-m+1,\ i-m}$ denotes the lateral current from node i-m+1 to i-m, the function f takes the form of equation 5 that describes the dependency of transistor current to its source and drain potential. I_{i-m} is the feedforward current to the output neuron.

When considering the case that with increasing m, the conductance of lateral becomes much lower than that of feedforward resistance, we have $I_{\text{i-m, i-m-1}} << I_{\text{i-m}}$, and $I_{\text{i-m+1, i-m}} \approx I_{\text{i-m}}$ as the boundary condition. Here, we consider a network of 1×9. The maximum lateral distance m to center input is 4. Initial value of $I_{\text{i-m}}$ was guessed and used to derive $V_{\text{i-m}}$, and $V_{\text{i-m+1}}$ according to equation 6-8. When the derived V_5 approaches the magnitude of input pulse (-0.5 V) with the error less than 0.5%, the iteration terminates.

In Fig. S9b, the derived negative current at each node were displayed as relative inhibitory weight (W_{I}) compared to the direct excitatory signal transmission (W_{E}=1 at input i). By adopting the square shaped input waveform with identical pulse width for positive and negative parts, the overall weight is simply calculated as $W_{SUM} = W_E + W_I$), which has the typical ON-center/OFF-surround characteristic. Since the weight drops rapidly with increasing distance >2 to input pixel, the function of lateral network could be approximated using a 3×3 convolution kernel displayed in Fig. S9c.

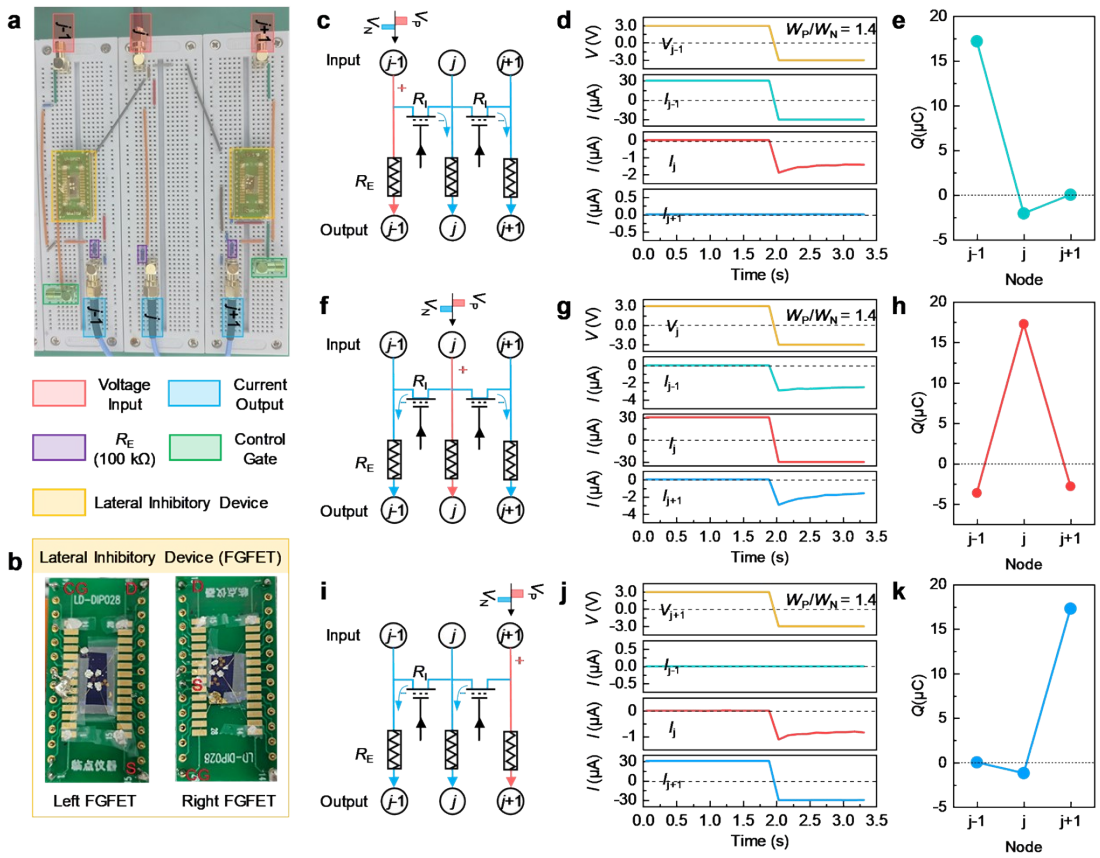

Fig. S10. Demonstration of lateral inhibitory weight using a 1x3 array, which has 2 lateral coupling FGFET. (a), (b) Image of the prepared circuit and the two FGFET transistor used in measurement. (c)-(e) The circuit configuration, monitored real-time current at output layer, and resulted effective weight (represented as integrated charge Q) when current is fed from pixel *j*-1. (f)-(h) and (i)-(k) displays the cases when spike waveform was input at pixel *j* and *j*+1, respectively. The lateral inhibitory weight decays rapidly with distance. Lateral coupling among nearest neighboring pixels can be reasonably considered to approximate the function of network.

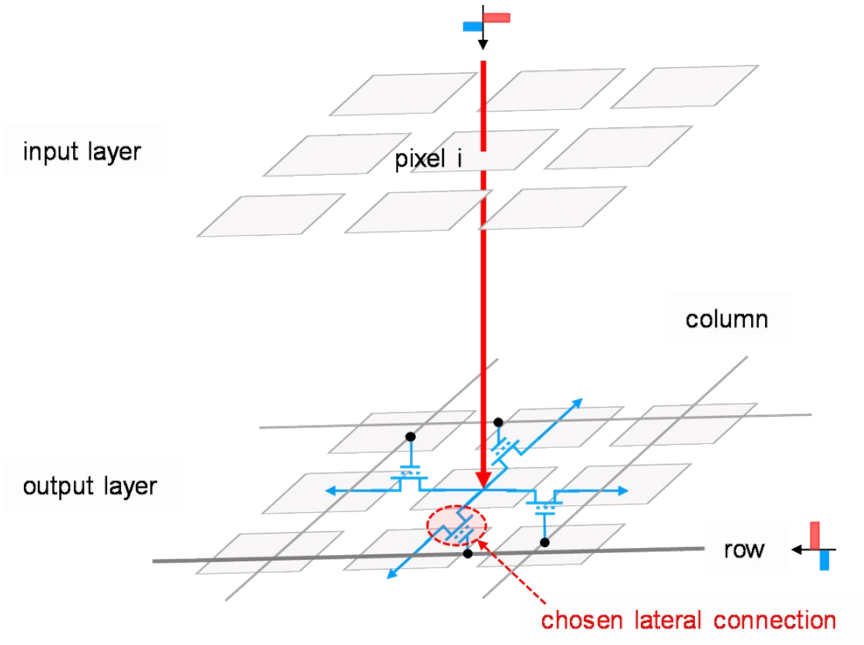

Fig. S11. The proposed strategy for local weight update. To modify the lateral coupling weight to local position i, a pre-spike can be applied at the input layer at the pixel-i, the specified lateral coupling transistor can be chosen by applying a post-spike from the row or column gate. By changing the time interval, the weight can be updated using the STDP rule. Alternatively, the addressing can be made by applying a program pulse to input pixel i, together with a selection gate bias with opposite polarity from column or row terminals. The number of program pulses can be changed to update the lateral inhibitory weight using a write-verify regime.