Supplementary Materials

Multifunctional UV Photodetect-memristors Based on Area Selective Fabricated Ga$_2$S$_3$/graphene/GaN Van der Waals Heterojunctions

Zhengliang Lin,[a] Junrui Chen,[a] Zhuohang Zheng,[b] Quanguang Lai,[a] Zhiqi Liu,[a] Liwei Liu[a], Jiaying Xiao[a] and Wenliang Wang[a, *]

[a]School of Materials Scicence and Engineering, South China University of Technology, Guangzhou, 510640, China

[b]School of Computer Science and Engineering, South China University of Technology, Guangzhou, 510640 China

*Corresponding author email: wenliangwang@scut.edu.cn
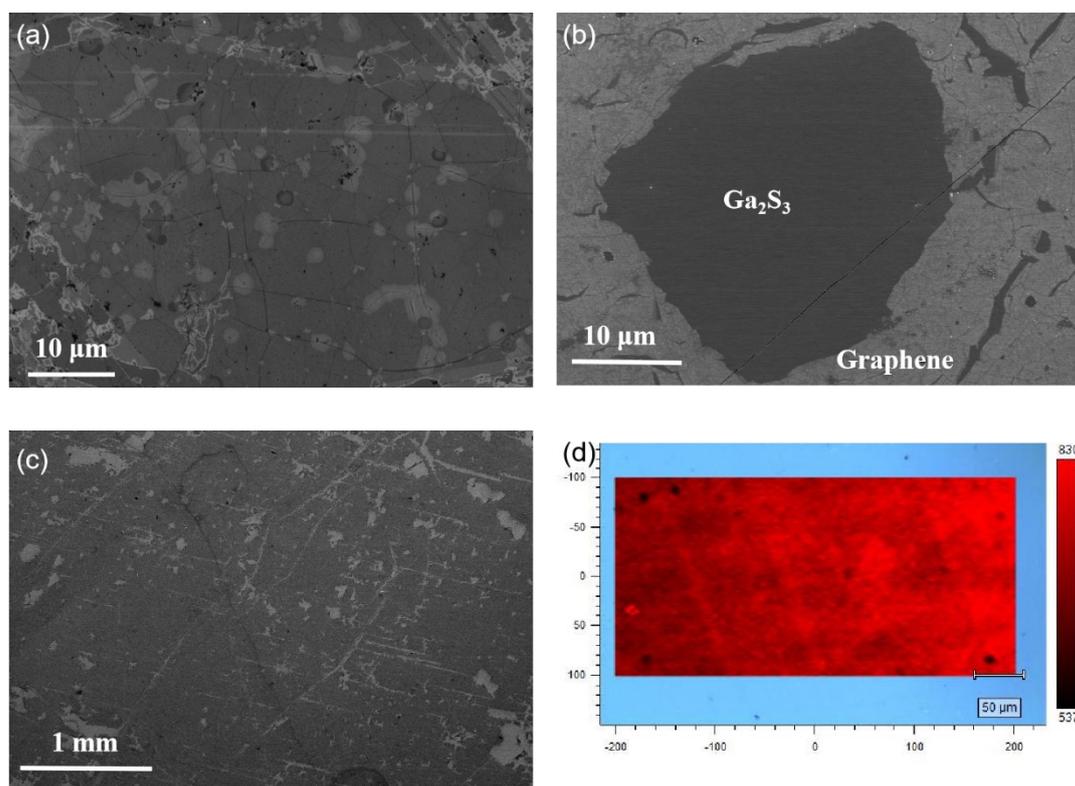
Fig. S1. SEM images of (a) high- and (c) low-magnification after area selective deposition (ASD). (b) SEM image of parameter modified Ga$_2$S$_3$ grown on graphene. (d) Raman mapping of graphene region after ASD.

Area Selective Deposition Simulation

Area selective deposition (ASD) is a very valuable and practical way to fabricate chips and has drawn lots of attention. This work realized high quality ASD by introducing

1

graphene interlayer, which directly skip the procedure of mask and lithography to reduce the total cost. In this section, the mechanism of ASD was studied systematically. Previous reports indicate that the process of growing 2D crystals on another one with similar lattice structure, where graphene and Ga_2S_3 are both hexagonal in this case, is more likely to be reproducible, energetically favorable, seamless stacking and cross growth[1]. While on the surface of GaN material, due to the existence of unsaturated bonds, the epitaxial mode is mostly island combination mode, and the precursor tends to cluster together and desorb at a low precursor concentration, making it difficult to form a continuous two-dimensional film[2]. Therefore, the thermaldynamic stability difference of substrates is considered to be a key parameter that leads to high quality ASD.

To further reveal the estimation above, density functional theory (DFT) was firstly applied to understand the thermodynamic mechanism of adsorption of Ga_2S_3 precursor atoms on GaN and graphene/GaN substrates. Here, the adsorption energy $\boldsymbol{E_{adsorption}}$ represents the preference of precursor atoms been adsorbed when passing through the substrates and is calculated as following[3]:

$$E_{adsorption} = \frac{(E_{subtrate} + nE_{adatom} - E_{total})}{n}$$

(1)

where $E_{subtrate}$ is the energy of the substrate, E_{adatom} and **n** are the energy and number of the adatom, respectively, and E_{total} is the energy of the surface with adatoms adsorbed onto it. Generally, a positive $\boldsymbol{E_{adsorption}}$ with higher value indicates a more stable adsorption model[4]. Here, two substrate models were established, respectively. The first one is the GaN substrate model with Ga face exposed in the air according to the experiments. Inspired by previous report[5], four adsorption sites were put into consideration(S_{top}, S_{bridge}, S_{fcc}, S_{hcp}). Similarly, the second model contains graphene/GaN substrate with four adsorption sites available(S_{mid}, S_{top}, S_{bridge}, S_{hcp})[6]. Additionally, the adsorption models are constructed with the coverage of 0.25, which has been widely accepted[3, 7]. The absorption sites on GaN and graphene are shown in

Fig. S2(a). In this simulation, Ga atom and S atom were used as absorption precursors according to previous report[8]. The calculated resulted of absorption energy of different precursors in different substrates are shown in Fig. 2(b) and (c), respectively. As shown, the $E_{adsorption}$ of two precursors on graphene/GaN substrate are relatively higher when comparing with GaN substrate in all absorption points, indicating a grater thermaldynamic stability. That means precursors are more likely to be absorbed in graphene/GaN substrate. Therefore, by controlling the growing condition in an appropriate interval where Ga_2S_3 2D material can be well formed on graphene while precursors tend to be desorbed on GaN, high quality ASD can be realized.
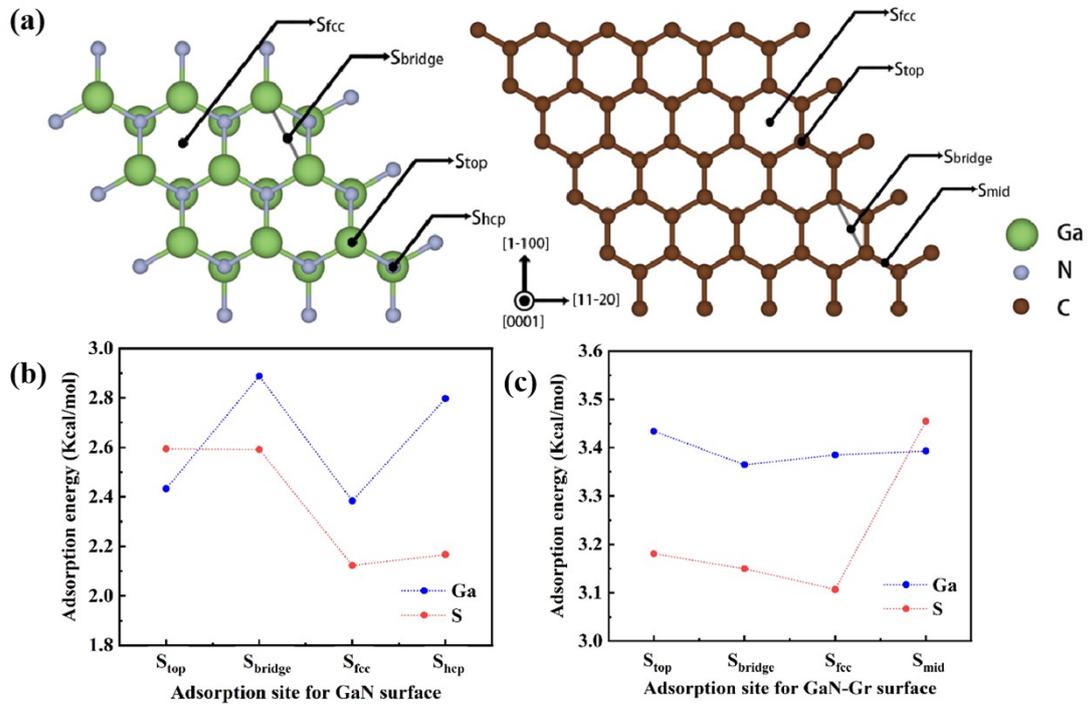
Fig. S2 (a) Schematic illustration of top views for GaN (left) and graphene (right) surfaces. Calculated adsorption energies for (b) GaN and (c) GaN/graphene surfaces with each Ga and H_2S precursors on high symmetric site.

What's more, the surface potential should also be put into consideration since it's of great importance to the absorption preference and distribution of precursor on substrates[9]. Referring to the principle of Kelvin scan probe in detecting the surface potential, the surface potential of two substrates is calculated as following[9]:

$$E_j = \sum_{i=1}^{n} 4\varepsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right)$$

(2)

here, σ and r_0 are separately used to describe the interaction distance of atoms pair in terms of the lowest absolute value Van der Waals potential or mole potential, which are different in their own cases. Similarly, ε represents the depth of potential trap in Van der Waals potential, respectively. The results of different precursor on different substrates were presented as heat map figures in Fig. S3(a)-(d). If the half of the depth of the potential valley was considered as the depth of the potential trap, it can be seen that the depth of both Ga and S precursors on the potential trap of graphene are around 10^{-22} J per atom, while that of GaN are around 10^{-23} J per atom, indicating that graphene has a higher surface potential than GaN. Moreover, as we can see, the unique structure of 2D material (graphene) provide a wider and deeper potential trap than 3D material (GaN), and it's believed that when the area of the potential trap is larger and with a more regular boundary, the atoms are easier to settle and jump from one potential trap to another, rather than escape from the substarte, which also results in an easier tendency of capturing atoms. Therefore, in terms of the surface potential, graphene has a higher tendency to absorb precursor atom than GaN.
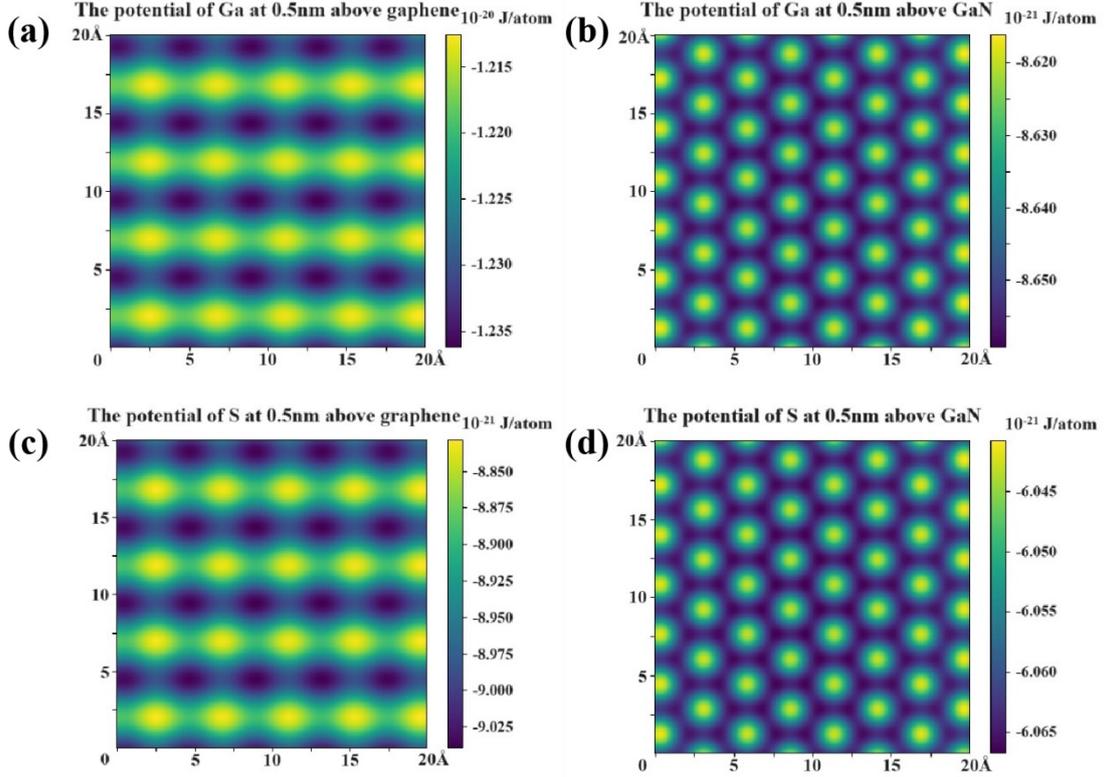
Fig. S3. Surface potential heat maps of Ga atoms on (a) graphene and (b) GaN and S atoms on (c) graphene and (d) GaN

Apart from that, the different growing modes on two substrates may also attribute to this amazing phenomenon, which is highly related to the surface tensions of the substrate, the growing film, and their interface, in another word, the Bauer criterion[10]:

$$\Delta\gamma = \gamma_f + \gamma_{s/f} - \gamma_s \qquad (3)$$

where γ_s, γ_f and $\gamma_{s/f}$ are the surface tensions of the substrate, the growing film and their interface, respectively. Depending on the $\Delta\gamma$ value, two possible growth modes can be distinguished: layer-by-layer growth or Frank-van der Merwe (FM) growth ($\Delta\gamma \leq 0$) and island growth or Volmer-Weber (VW) growth ($\Delta\gamma > 0$). When the $\Delta\gamma$ value changes its sign during the film growth, the third mechanism, layer-by-layer/island growth is possible, which is known as the Stranski-Krastanow (SK) mode. Since GaN has dangling bond, which greatly increase the tension between the Ga_2S_3/GaN interface, while no tension between Ga_2S_3/graphene interface since it's bonded by van der Waals force. Therefore, it's believed that precursor has formed VW growth on GaN and FM growth on graphene, where the former one has been suppressed under the growing

condition due to lower thermodynamic stability, leading to ASD. Apart from that, TEM image in Fig. 1(i) reveals that the Ga_2S_3 film grown on graphene has a very flat and uniform feature, which indicates that the transition of growth modes from FM to VM may occur when the it reaches the critical thickness since the tension of substrate is gradually changed from substrate to Ga_2S_3 itself as the Ga_2S_3 film being formed. Similarly, the VM growth mode, which often leads to the formation of discrete bulk overgrown particles known as "adlayer" reported previously[8, 11], is suppressed as well, leading to high quality 2D film.

To further investigate the factors influencing the selectivity, molecular dynamic (MD) simulations were conducted. MD simulation is a well-developed method to investigate the behavior of particles in molecular sizes. It uses the potential energy functions to describe the interaction forces of particles and integrates the Newtonian equation of motion to get the trajectories of particles[12]. The trajectories of these particles can reflect a macroscopic behavior. Some studies have used molecular dynamics to investigate the interaction on interfaces, which explained and predicted the results of experiments[13-15]. In the experimental part, a complete chemical vapor deposition (CVD) method included deposition, island growth, and crystallization. The selectivity of crystallization would be verified if the deposition rates on each substrate were completely different[16]. Compared with mainstream simulation software such as LAMMPS, GROMACS, and Material Studio, the self-programming MD simulation program based on Python offers a more flexible environment for users to adjust the algorithm according to the conditions of the experiment and was used to confirm the result of the experiment and show the selectivity of different substrates of gas atom deposition.

Considering the balance of running speed, reaction time, and amount of gas atoms, an ensemble was built with the dimensions of 120.4 nm × 120.4 nm × 80.3 nm, which could contain 144 atoms in total. The thickness of both graphene and GaN substrate were set to be over 1.5 nm and their component atoms were fixed so as to reduce the complexity. Ga, H, and S atoms were introduced into the NVT ensemble with the ratio

of 1：2：1 to create the desire gaseous form of hydrogen sulfide (H_2S) and free state Ga atoms, where the configuration of the H_2S was calculated by geometric optimization *via* DFT. Then, DFT calculation was also used to determine the interaction strength between atoms. Several models were built according to different interaction pairs and calculated in regard of their potential energy at different distances. For non-bonding interaction, it was described with Van der Waals potential while bonding interaction was in turn described with Morse potential as formula[12]:

$$E_{non-bonding} = \sum_{\substack{i=1 \\ i \neq j}}^{n} 4\varepsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right)$$

(4)

$$E_{bonding} = \sum_{\substack{i=1 \\ i \neq j}}^{n} D_0\left(e^{K(r_{ij}-r_0)} - 1\right)^2$$

(5)

Among them, σ and r_0 are separately used to describe the interaction distance of atoms pair in terms of the lowest absolute value Van der Waals potential or mole potential, which are different in their own cases. Similarly, ε and D_0 represent the depth of potential trap in Van der Waals and Moore potentials themselves, respectively. K is called the depths of quantum well and k is constant. It's worth mentioning that formula 13 is mostly the same as formula 2 but with different description object, where formula 13 is for the single atom pair while formula 2 is for the whole system. The DFT calculation results summary is presented in Table S1.

Non-bonding interaction pairs	σ(J)	ε(m)
Ga-H	3.628×10^{-22}	3.185×10^{-10}
Ga-S	3.767×10^{-21}	4.161×10^{-10}
Ga-Ga	3.074×10^{-21}	4.3558×10^{-10}
Ga-N	4.891×10^{-21}	3.5868×10^{-10}
Ga-C	9.656×10^{-21}	3.8429×10^{-10}
H-C	1.529×10^{-22}	3.3106×10^{-10}
H-N	4.349×10^{-21}	3.4264×10^{-10}
S-C	8.950×10^{-21}	3.6825×10^{-10}
S-N	1.129×10^{-22}	3.0605×10^{-10}

Bonding interaction pair	$D_{0(J)}$	$K_{(m^{-1})}$	$r_{0(m)}$
Ga-S	5.139×10^{-19}	-0.826×10^{10}	2.240×10^{-10}
H-S	5.163×10^{-19}	-1.925×10^{10}	1.326×10^{-10}

Table S1. Summary of particle interactions calculated from DFT.

To control the temperature of the ensemble, a local Berendsen thermostat was introduced onto the top and bottom of the ensemble, and the thermal exchange coefficients were set according to the thermal transmission coefficient of graphene and GaN[17-19]. The thickness of the thermostat was designed to be 5 Å, at which point the heat transformation is the strongest so as to realize the particle collision. To ensure the accuracy and enhance the efficiency of the simulation, a cutoff distance was designed to be 15 Å. The calculation time step and output time step were set to 0.001 ps and 0.1 ps. The total simulation time would vary from different intentions.

To show the selectivity of gas atoms deposition, it is needed to developed a descriptor to count up the deposition atoms on the substrate. In this work, considering the vibration of atoms on the substrate, the atom located at a vertical distance smaller than 8 Å between it and the substrate will be regarded as being deposited. Additionally, since the experiment had shown that the thin film only contained Ga and S, only these two elements would be counted up. Thus, the deposition rate should be defined as:

$$\theta = \frac{d_1 + d_2}{t_1 + t_2}$$

(6)

Where d_1 and d_2 refers to the amounts of deposited Ga and S atoms, t_1 and t_2 refers to the total amounts of Ga and S atoms inside the ensemble. Based on these preconditions, deposition statistics could be conducted and shown with the deposition rate versus times. For more information about the model setting, please referring to the method part. Firstly, a 1500 ps simulation was carried out in terms of the substrates difference to reveal the selectivity qualitatively and the real simulation images were recorded and displayed in Fig. S4(a)-(f). As shown, both of the deposition rate curves increase with the advancement of time and eventually reach their saturation states. The precursor particles show obvious selectivity of deposition, where the maximum deposition rate of graphene was almost 10 times higher than that in GaN. These results directly reveal the deposition preference of precursor on graphene compared to GaN.
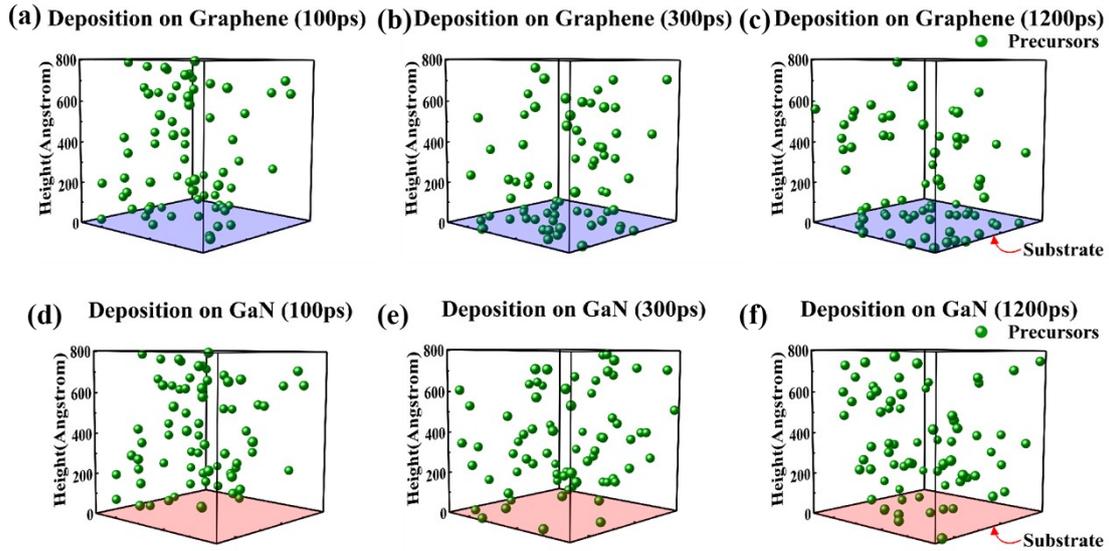
Fig. S4. Pictures of precursor absorption model on GaN recorded at (a) 100 ps, (b) 300 ps, (c)1200 ps and that on graphene recorded at (d) 100 ps, (e) 300 ps, (f)1200 ps.

To conclude these phenomena the relationship curves of precursor deposition rate and time on different substrates are shown in Fig. S5. Comparing to graphene, which has a saturated deposition rate of around 0.5, the curve on GaN is going through a regular vibration between 0.05 and 0. This indicates the change of weak equilibrium between absorption and desorption due to the variation of precursor concentration on the GaN substrate.
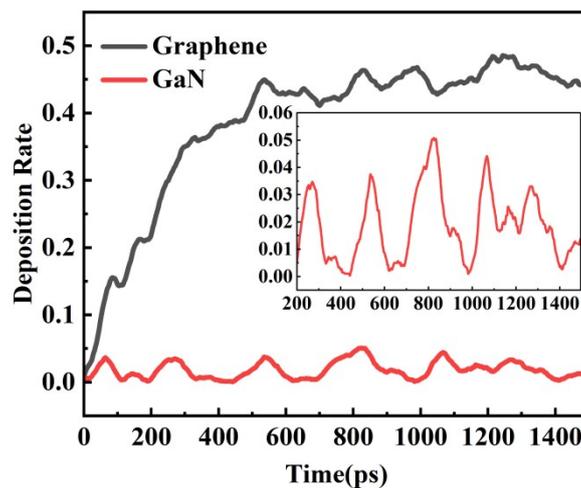
Fig. S5. The relationship between deposition rate and time in terms of different substrates. The inset figure is the enlarged picture of GaN.

The growing parameters could also affect the quality of ASD. Here, several conditions

were put into consideration. Firstly, the substrate temperature was put into consideration. Since the temperature of substrate will decrease as the distance between the heat source increase, a temperature coefficient k was defined to describe the relationship between them mathematically:

$$T_s = kT_g \tag{7}$$

where T_s is the temperature of the substrate, and T_g is the setting temperature. In order to have a comprehensive understanding on how the modification of substrate temperature affect the performance of selective deposition, the values of k were set to be between 0.4 and 0.9 with the step size of 0.1 to conducted the similar simulation correspondingly on different substrate. Learned from the above case, a distinct difference could be observed in a relatively short period. Thus, several 200 ps based simulations could be feasible to study the temperature-dependent deposition behavior as well as reducing the total computation time and were displayed as Fig. S6 (a)-(f).
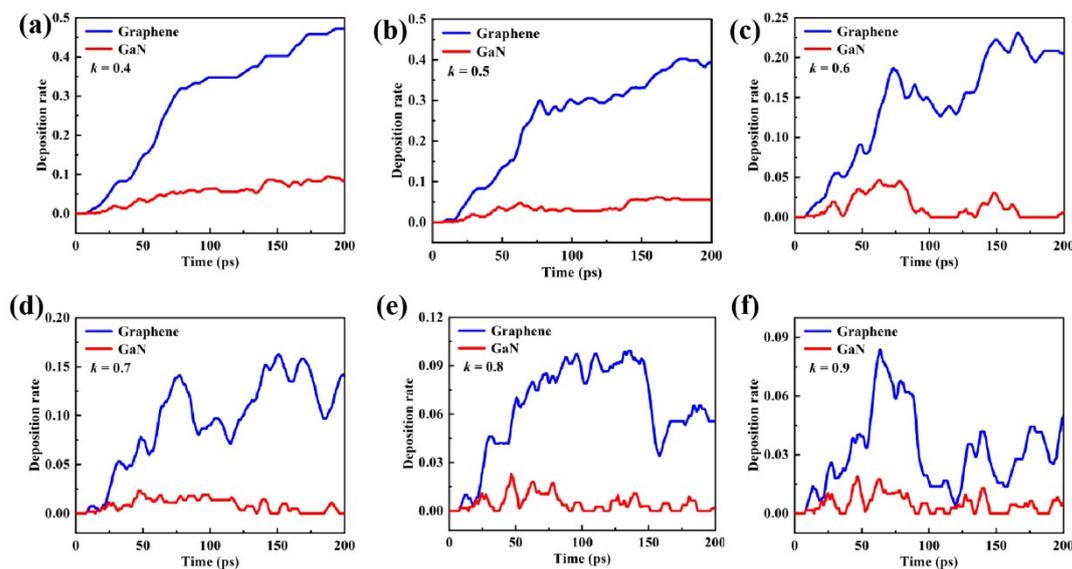
Fig. S6 Comparation of deposition rate verse time in terms of GaN and graphene when temperature coefficient k is (a) 0.4, (b) 0.5, (c) 0.6, (d) 0.7, (e) 0.8 and (f) 0.9.

Similarly, simulation at different pressure with a pressure index i, which is the exponent of pressure, was put into consideration:

$$P_i = P_0{}^i \tag{8}$$

Where P_i and P_0 represent the pressure of setting one and default one. i values were taken from 0.8 to 1.3 with the step size of 0.1. The results of similar description from above were plot on Fig. S7(a)-(f).
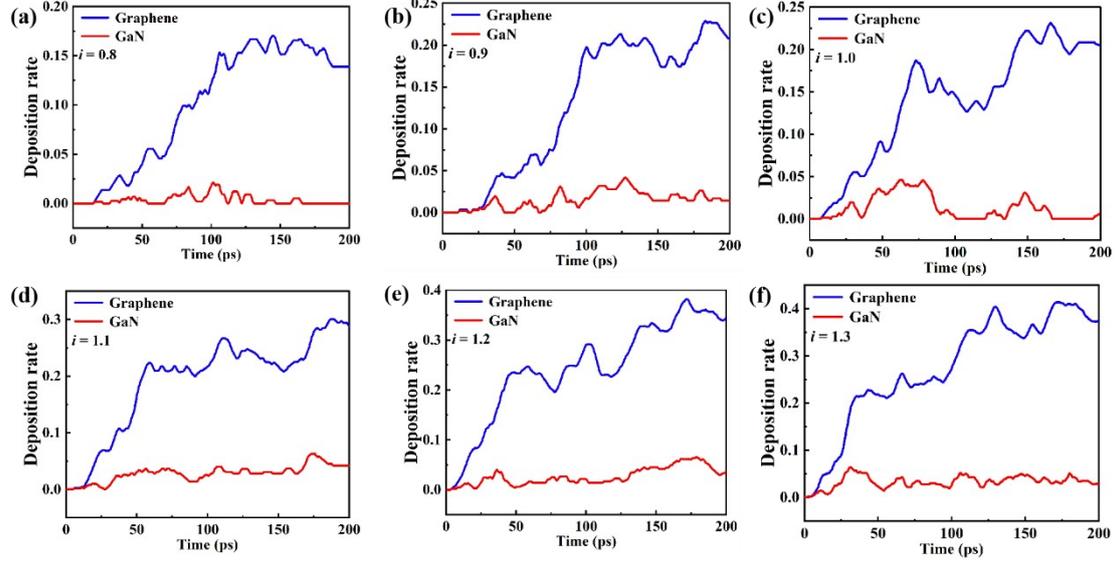
Fig. S7 Comparation of deposition rate verse time in terms of GaN and graphene when pressure index i is (a) 0.8, (b) 0.9, (c) 1, (d) 1.1, (e) 1.2 and (f) 1.3.

Based on these, quantitative selectivity S was introduced and defined as:

$$S = \frac{\theta_1 - \theta_2}{\theta_1} \tag{9}$$

where θ_1 refers to the deposition rate of graphene, and θ_2 refers to the deposition rate of GaN. What's more, to evaluate the kinetic behavior, the deposition rate was averaged with time to get a more representative value, which is called expected deposition rate.

$$\overline{\theta}_{\frac{a+b}{2}} = \frac{\sum\limits_{a < t < b} \theta_i \Delta t}{t_b - t_a} \tag{10}$$

Where θ_i is the instantaneous deposition rate in time i, Δt, t_b, t_a is the length of time, and the initial time and ending time.

The relationships of the k value, $\overline{\theta}_{100ps}$, and S are shown in Fig.S6 (a). The result

confirms that the deposition kinetic was negatively related to the temperature for both substrates. It shows that the relative difference was higher than 0.8 for almost all the temperatures and even higher than 0.9 when k is between 0.6 to 0.8. This is because as the deposition rate of both goes down together with increasing k, there is a zone that the deposition rate of graphene falls slower than the GaN, resulting in a relatively higher selectivity in deposition. The results indicates that selectivity exists in a wide range of temperatures, setting the foundation of industrialization. The relationships between i, $\bar{\theta}_{100ps}$, and S is shown in Fig. S6 (b), which presents the trend that the speed of deposition is positively related to the pressure, and the selectivity is negatively related. A slower deposition may result in a high selectivity, but reaction at such a low speed is not suitable for massive production. While a faster deposition would lead to a relatively low selectivity.
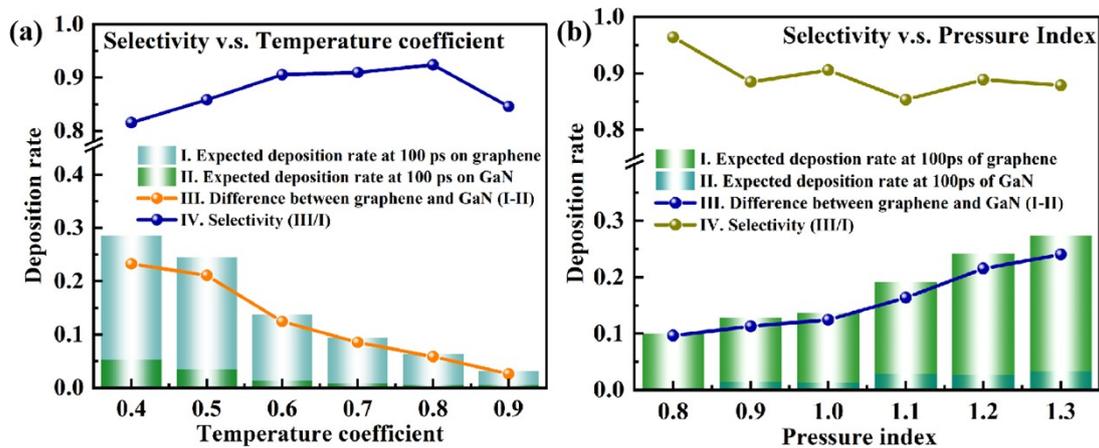
Fig. S8 The relationship of deposition rates/ related mathematic relation and (a) temperature coefficient/ (b) pressure index in terms of different substrates.

The above result implies that the external conditions have a certain influence on the deposition kinetic, but was not as evident as that of substrate since the selectivity are always higher than 0.8 in all cases. On the whole, the simulation results show that the selectivity exist in a wide range of conventional working condition of CVD process which provides a fundamental of mass production of ASD based Van der Waals heterojunction devices.
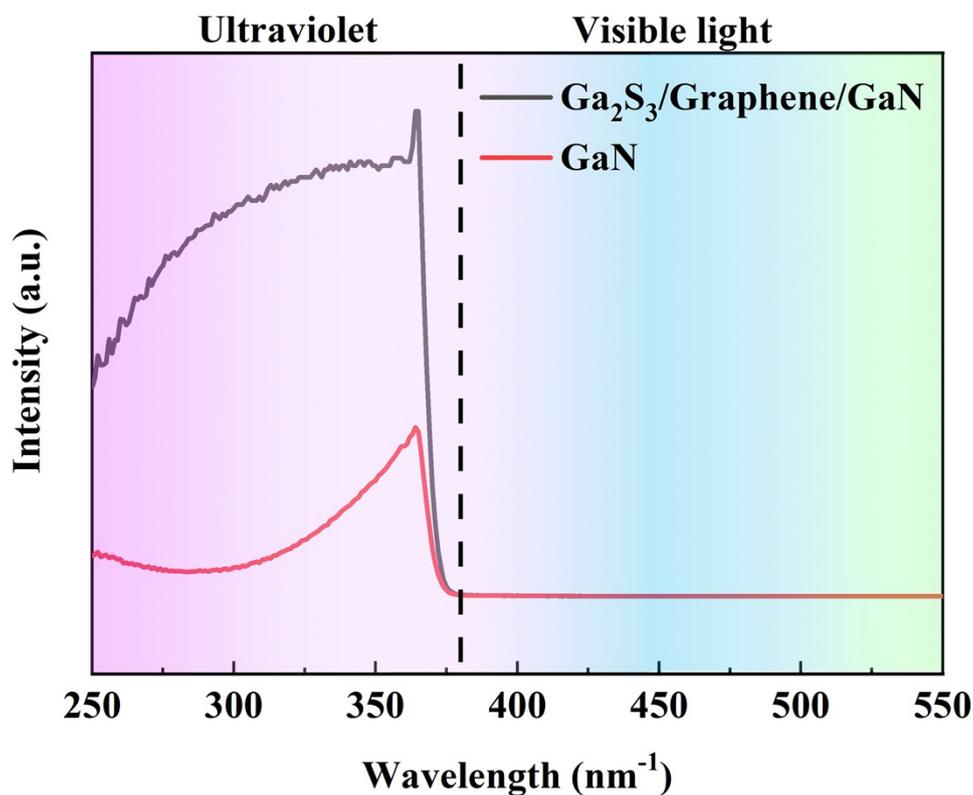
Fig. S9. Photoluminescence excitation (PLE) of bare GaN and heterojunction.
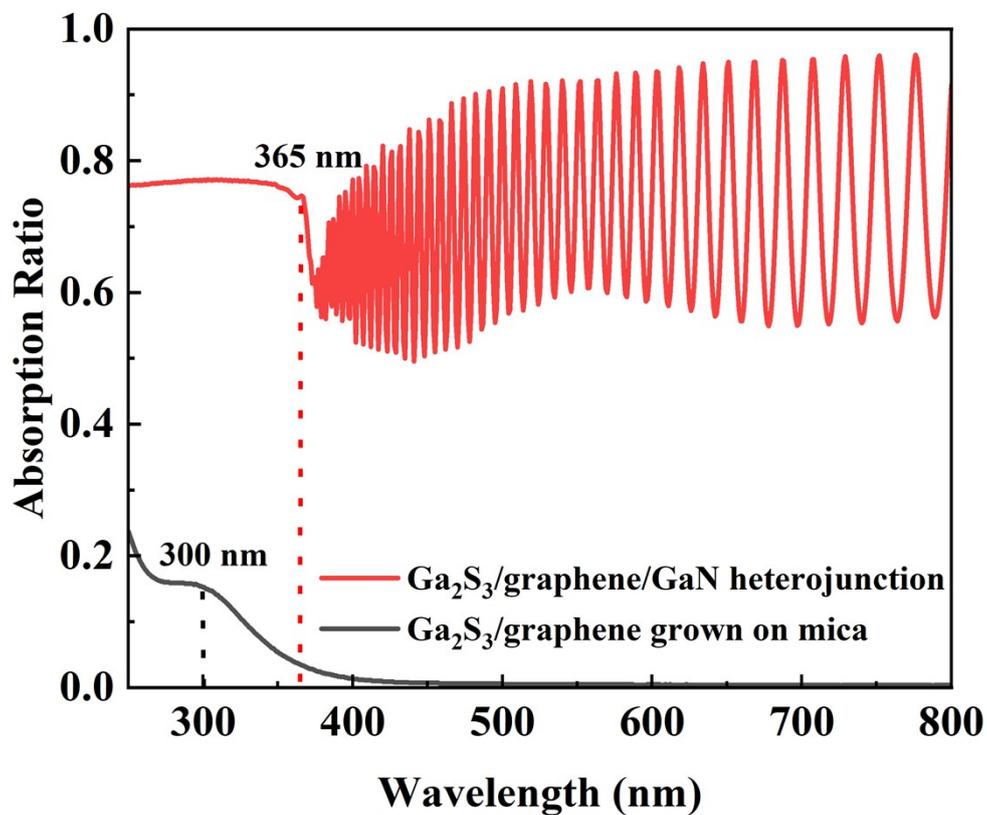
Fig. S10. UV-vis absorption spectra of Ga_2S_3/graphene/GaN heterojunction and Ga_2S_3/graphene grown on mica.
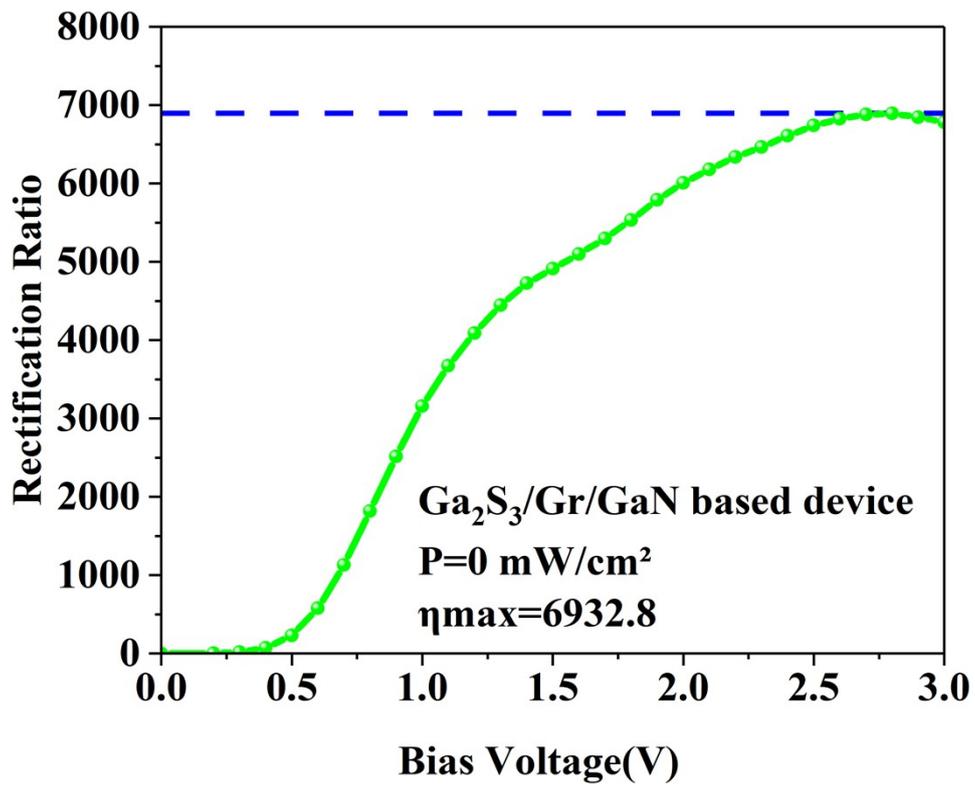
Fig. S11. The rectification performance of Ga_2S_3/graphene/GaN-based device
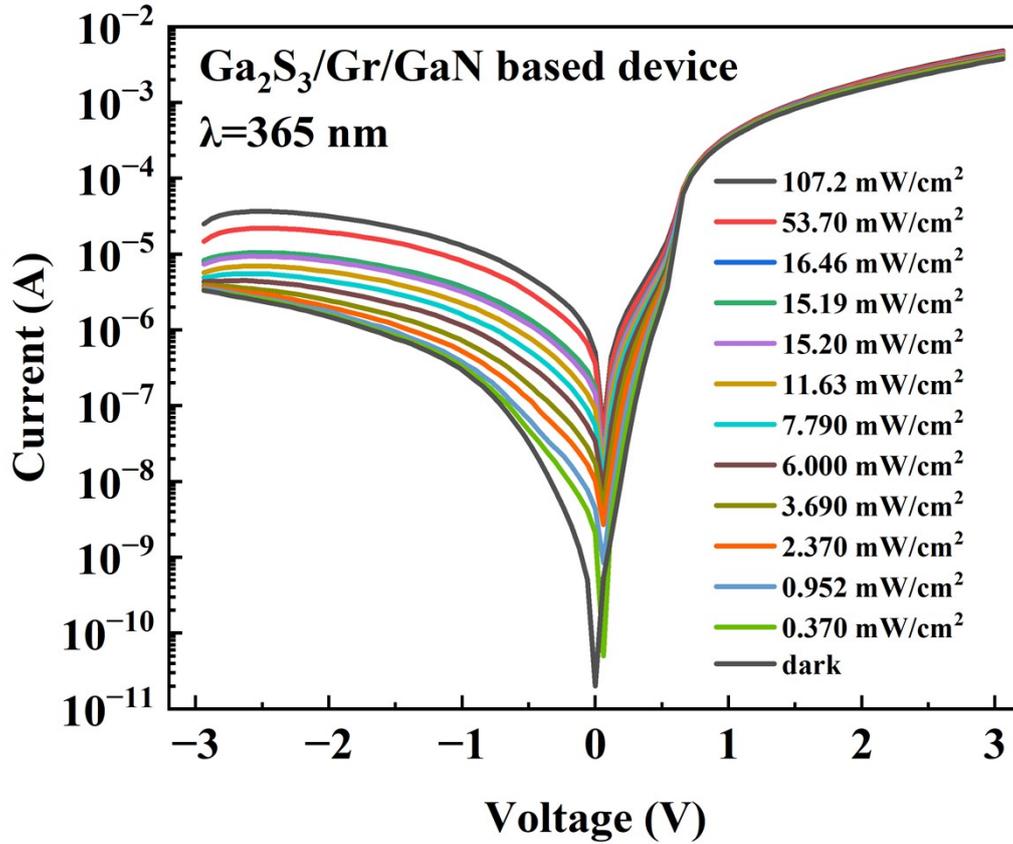
Fig. S12. *I–V* curves of device as photodetector in the dark and under UV illumination.
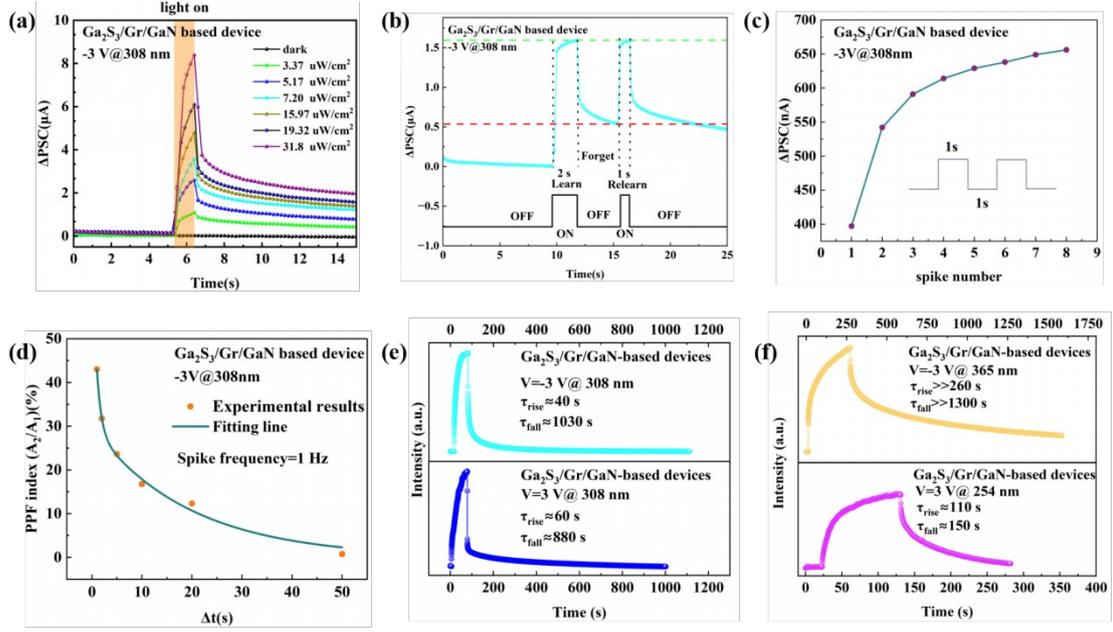
Fig. S13. Performance of device as optomemristor. Photoresponse under 308 nm UV light (a) with various light intensities. (b)The Learn-Forget-Relearn curve under UV illumination. (c) The difference of postsynapse current (ΔPSC) plotted as a function of the spike number (d) Paired-pulse facilitation (PPF) index of as-fabricated device. (e) *I-T* curves under -3 V@308 nm and 3 V@308 nm. (f)*I-T* curves under -3 V@365 nm and 3 V@254 nm. Here, t_1 and t_2 is 0.90 s and 19.47 s under the impulse frequency of 1 Hz according to PPF index in Fig. S9 (d)[20], which are also consistent of reported optomemristor[21-23]. It should also be noticed that in these cases, the proportion of slow decay in the decrease current after the removal of illumination is larger than before (Fig. 3(d)), which is because higher illumination intensity increases the ratio of carriers blocked by barriers compared to that collected by electrodes so that the optomemristor plays a more important role.

Device Performance Definition

Firstly, the responsivity (***R***), also known as the photodiode current per unit of incident light per unit of power, can be calculated using the following formula:[24]

$$R = \frac{I - I_{dark}}{PA}$$

(11)

Where ***I*** and ***I_{dark}*** represent currents measured under light and dark conditions, respectively; ***P*** is the power intensity of incident light and a is the illumination effective

area of PD (0.125 mm^2) . **D*** is the detectivity, defined as the index of the detector's ability to detect the minimum signal, also known as the detector sensitivity, which can be calculated using the following formula:[25]

$$D^* = R\left(\frac{A}{2eI_{dark}}\right)^{\frac{1}{2}}$$

(12)

Where every symbol is the same as above except **e** represents the electron charge. External quantum efficiency (**EQE**), is considered to be a direct expression of the relationship between the induced photon and the carrier that ultimately produces the response. This is calculated as follows:

$$EQE(\%) = \frac{1240R}{\lambda}100$$

(13)

Here, λ represents the wavelength of incident light. Lastly, linear dynamic range (**LDR**) describes the range of linear response of PDs to the light intensity. It can be calculated by[26]:

$$LDR = 10log(\frac{P_{sat}R_{365\,nm}}{\sqrt{2eBI_{dark}}})$$

(14)

where **P$_{sat}$** is the saturated light power when it deviates from the linear response (which is much larger than the light power at the density of 676 mW/cm^2), R$_{365nm}$ is the photoresponsivity, and **B** is the bandwidth, which is about 7.7×10^{-4} Hz (Details can be found in Fig. S13(f)).
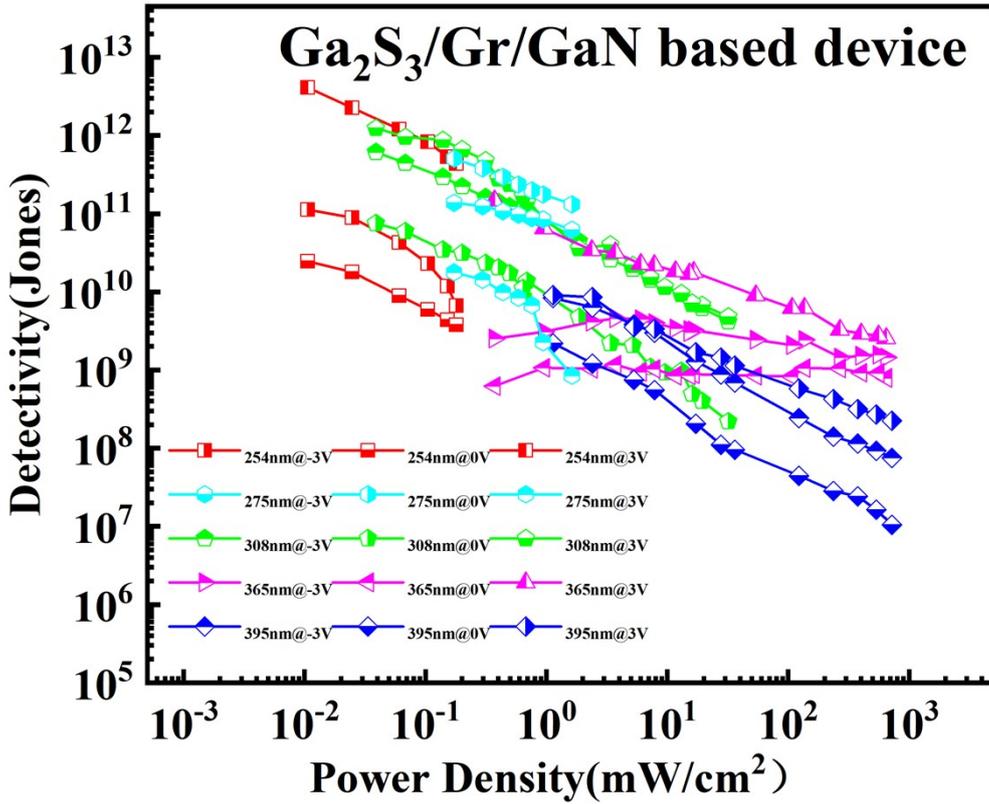
Fig. S14. The dependence of detectivity on light power density under various wavelengths of UV illumination and different bias modes.
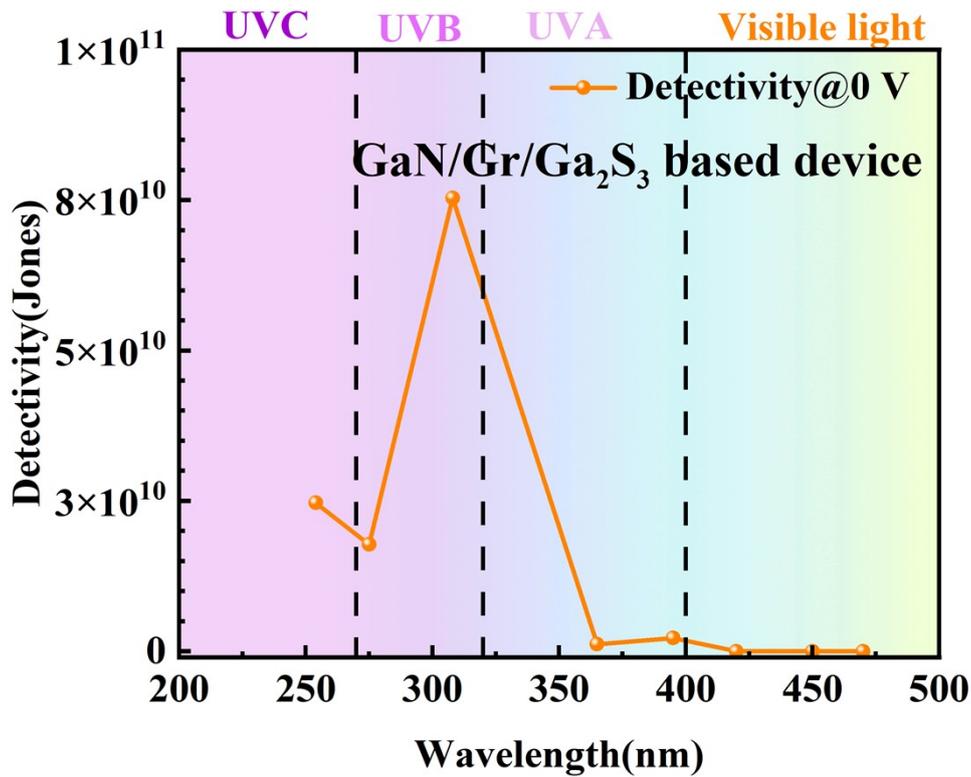
Fig. S15. The dependence of detectivity on light power density under various wavelengths of UV illumination and different bias modes.

Device Working Mechanism Revelation via *I-V* Fitting

In this work, the *I-V* curves with the bulk dominated mechanisms including Ohmic ($I \propto V$), space-charge-limited current (SCLC, $I \propto V^n$) and Poole-Frenkel emission (P-F, $\ln(I/V) \propto V^{1/2}$), and interface dominated mechanisms including Schottky or thermionic emission ($\ln(I) \propto V^{1/2}$) and Fowler-Nordheim tunneling (FNT, $\ln(I/V^2) \propto 1/V$)[27-30] were put into consideration. It was found that the P-F emission, SCLC and thermionic emission can fit the *I–V* curves well, showing the linear relationship at the relatively large voltage region. However, considering the almost defect free 2D functional layers in our device, the synaptic properties should be mainly determined by the interface effects, meaning that the thermionic emission plays a dominating role in the resistive variation process as shown in Fig. S16(a), the result is consistent with previous analysis that charge carriers transport are limited by the height of the interface barrier. If the carriers can obtain enough energy provided by thermal activation, they will overcome the potential barrier at the semiconductor/graphene interfaces[31, 32]. As the input photons increase, the photogenerated carriers gradually fill the potential well in graphene layer, resulting in a higher response current that depends on the increased free charge carrier concentration. Fig. S16(b) is the fitting result with the thermionic emission model of the device in dark. The linear relation of $\ln(I)$ versus $V^{1/2}$ in the plot shows that the device in dark has similar conduction mechanism compared with the device under light, which is consistent with the thermionic emission mechanism.
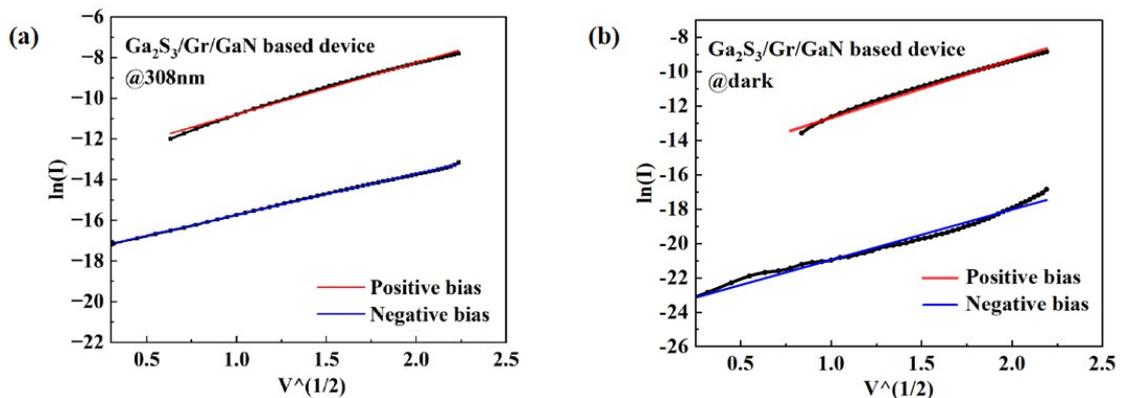
Fig. S16. Fitted *I-V* curves with thermionic emission conduction mechanism (a)with illumination under the wavelength of 308nm and (b)without illumination.

ESN-based Pattern Recognition Simulation

ESN is a variant of RNN, which is mainly used to process temporal data. A supervised learning algorithm, linear regression, is used by the network as the training method, instead of gradient-related and backpropagation, which avoids the problems of gradient exploding, heavy computational costs, local minima, etc. At the same time, it is conceptually simple, efficient training, easy implementation, etc. It consists of the input layer, the reservoir, and the output layer. The reservoir is the core of the ESN, and it is a pool of randomly interconnected nonlinear neurons that produce the reservoir states for feature extraction by mapping the input temporal data to the reservoir. The temporal data can be decomposed into a series of time series data according to the specific time steps, which makes the reservoir get a large number of time-dependent reservoir states.

The input neurons representing the input features are connected randomly and sparsely to the reservoir neurons, which can transform the input feature vector u[n] composed of time series data of the current time step n into the reservoir input vector *r'[n]* of the current time step *n* by the following equation[33]:

$$r^{'}[n] = g\left(W^{in}u[n] + w^{bi}\right) \qquad (15)$$

where W^{in} and W^{bi} are the input weight matrix and the bias weights of the input layer, and g is the activation function. Theoretical studies have suggested that the dimension N^{in} of the input feature vector should be much smaller than the connections K^{in} between each reservoir neuron and each input neuron [33], so the sparsity of the network is defined as the proportion of *K^{in}* in *N^{in}*.

The calculation of the reservoir state *r[n]* of the current time step n needs *r[n-1]* for making the reservoir neurons have recurrent connections and a great memory capacity, which allows the reservoir to capture the temporal correlations of the input temporal data. To quantitatively measure the memory capacity of the reservoir, the leakage rate *λ* of the reservoir neurons is defined, which denotes the leakage of the reservoir state at the previous time step. Then *r[n]* can be computed as follows [33]:

$$r[n] = (1-\lambda)r[n-1] + \lambda f\left(r^{'}[n] + W^{res}r[n-1]\right) \qquad (16)$$

where W^{res} is the reservoir weight matrix of the connections between the reservoir neurons, and f is also the activation function. The reservoir state collection matrix R can be obtained by the above two equations. During the training process, to minimize the output error between the output collection matrix Y and the target output collection matrix D, the output weight matrix W^{out} which connects the reservoir neurons to the output layer neurons is calculated by the methods of least square optimization in linear regression and generalized inverse matrix. The readout function is [33]:

$$W^{out} = \left(RR^T + \epsilon I\right)^{-1}\left(DR^T\right) \tag{17}$$

where ε is the regularization factor and I is the identity matrix. In fact, the equation comes from the ridge regression. The calculation formula of the output vector $y[n]$ at time step n is [33]:

$$y[n] = W^{out}r[n] \tag{18}$$

In addition, W^{in}, w^{bi}, and W^{res} are randomly generated and remain fixed when training, which is the key reason for the reduction of computational cost and training time.

When fitting the simulated memristor model, the experimental data was used under the experimental conditions of 0.1 Hz pulse frequency and 1 s pulse width. The curve fitting tool box in MATLAB was used to fit the increasing and decaying parts of the PSC in the selected experimental data with the function formula, respectively. The former one uses a custom polynomial function, while the latter one uses a custom exponential function with three terms to get the PSC activation function and the PSC decay rate function. When a dynamic memristor receives a series of pulses, the activation function and decay rate function can be used to calculate and simulate the PSC change process. In this way, the dynamic memristor can be regarded as a combination of the two functions, and produce corresponding responses to the input pulses. The optimized result of the activation and decay function are shown as following:

$$f_{up}(t) = 5017.78t^5 - 1518.14t^4 + 1720.75t^3 - 9092.24t^2 + 2408.55t + 0.01$$
$$(0s \le t \le 1s)$$
$$\tag{19}$$

$$f_{down}(t)$$
$$= 0.2266e^{-0.0122(t-1)} + 0.1106e^{-0.1821(t-1)} + 0.6408e^{-2.0062(t-1)} + 0.0221 \quad (t \geq 1s)$$

(20)

To verify the effectiveness of the fitting function, firstly, the functions were modified and utilized to get the experimental data of eight consecutive stimulus signals, where the results are shown in Fig. S17 (a). The comparison of the PSC between the fitting data and the experiment data shows that the simulated PSC curve can well describe the experimental PSC curve except some difference at the beginning. In order to explain this phenomenon from a deeper perspective, a comprehensive study of decay rate extracted from the leakage rate needs to be calculated. Here, the sum of the leakage rate and the decay rate is 1 and reason of why using decay rate to describe is because this is more alike to the decaying tendency of fitting curve. Fig. S17 (b) shows the combination of 8 decay rate curves extracted from the experiment and the simulation one. Recall from Fig. 3(f), the decay rate is highly related to the size of the PSC peak. The higher PSC peak is obtained with continuous pulses, which leads to a slower decline of the decay rate owing to the accumulation of light-induced carriers. That is to say, the effect of carrier leaking is smaller with higher carrier concentration. It is observed that the fitting curve is close to the experimental curves of the higher PSC peaks.
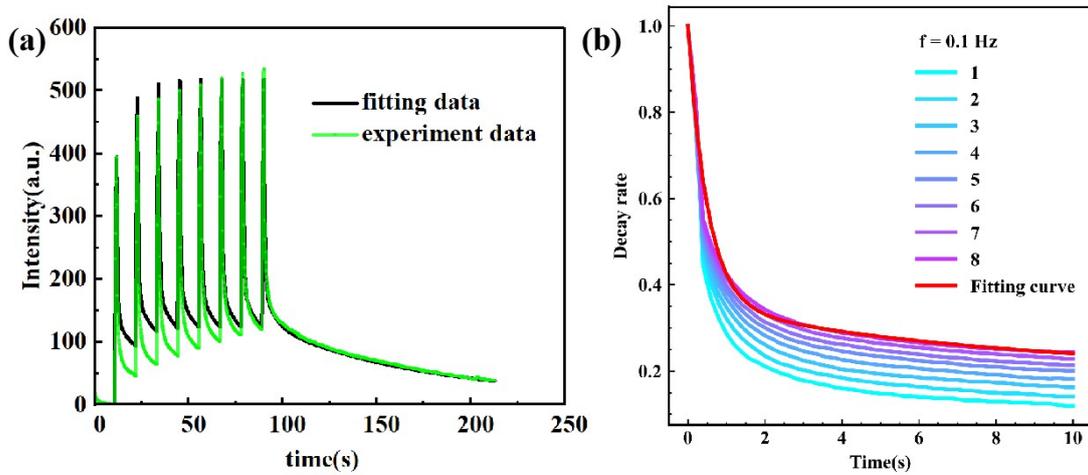
Fig. S17. (a) Comparison of the PSC between the fitting data and the experiment data of dynamic memristors. (b) Comparison of decay rate curve with the time changing from 0 s to 10 s between the fitting data and the experiment data.

To better describe the performance of the fitting curve, Table S2 summarizes the evaluation metrics of the fitting effect, which further confirms the fitting effect. From the table, the R^2 of 0.92 indicates the fitting data has high fitting degree. The MAE of 14.15 and the MRE of 15.129% indicate the influence of fitting error is small. The RMSE of 23.704 is close to the MAE, which means there are no outliers' data. The MAPE of 10.6% indicates the margin of error for the fitting data is small. Considering that the calculation of the decay rate is extremely complex and increases the fitting cost, the fitting curve is adequate enough to represent the experimental results.

In terms of the optimized parameter training process, 80% of the UCI handwritten digits dataset (1,797 test samples, 8×8 pixels, gray level range of 0-16) was divided into a training set while the other 20% was divided into a test set. Apart from that, the time step interval, time step and the number of selected pixels per row K_{in} were put into consideration as controlling parameter and the number of dynamic memristors in the reservoir is fixed as 400. The effect of these parameters to the recognition accuracy are shown both as figure (Fig. S18) and table (Table S3). The optimal network parameters were determined to get the best classification results, where the binarization threshold value of the grayscale images was 4, the pulse width was 1 s, the number of selected pixels in the same row was 2, the width of time frame and time step interval were 4 s and 2 s and the highest recognition accuracy is 94.2 %.

Parameters	R-Squared	MAE	RMSE	MRE	MAPE
Values	0.920	14.150	23.704	15.129	0.106

Table S2. The evaluation metrics of the fitting effect for fitting data (round to 3 decimal places), which includes coefficient of determination (R-Squared), mean absolute error (MAE), root mean square error (RMSE), mean relative error (MRE) and mean absolute percentage error (MAPE)
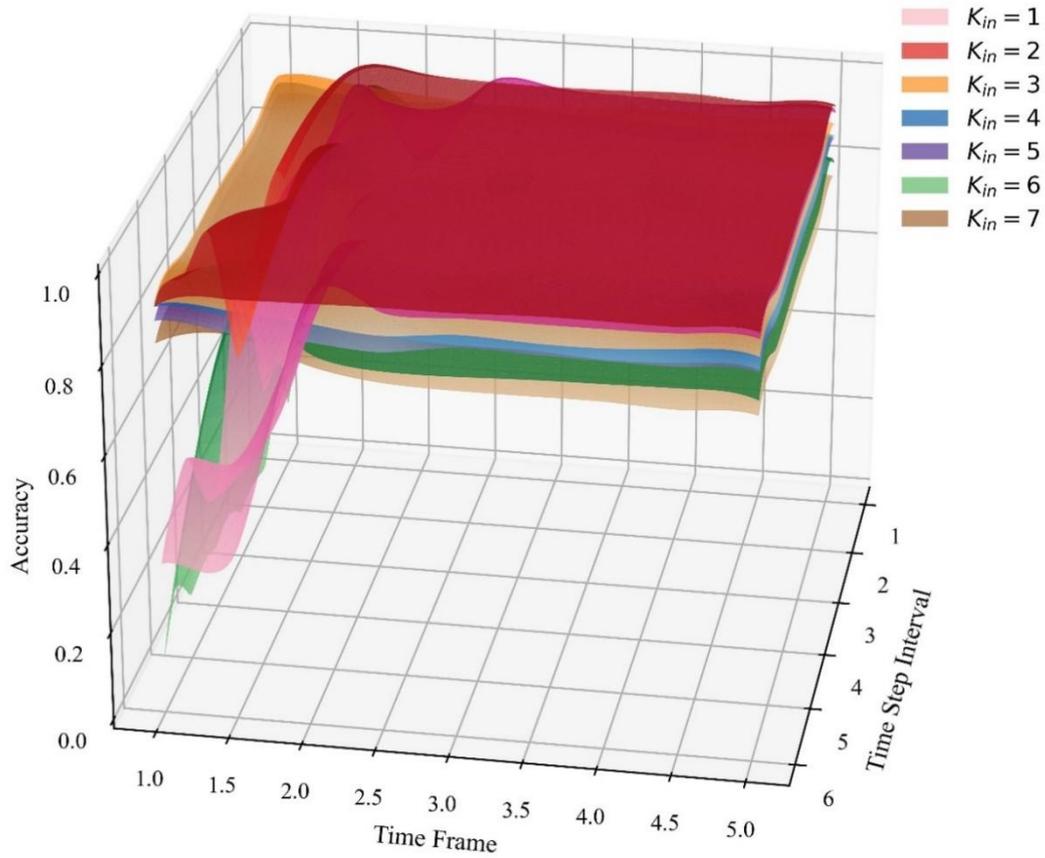
Fig. S18 The recognition accuracy changing under different parameter combinations. The time frame is from 1.0 s to 5.0 s in step of 0.5 s, the time step interval is from 1 s to 6 s in step of 1s, and $\mathbf{K^{in}}$ is from 1 to 7. Every $\mathbf{K^{in}}$ corresponds to a surface of recognition accuracy.

Time frame/(s)	K^{in}	Interval time/(s)						Time frame/(s)	K^{in}	Interval time/(s)					
		1	2	3	4	5	6			1	2	3	4	5	6
1.0		0.156	0.264	0.614	0.253	0.447	0.322	1.0		0.872	0.867	0.867	0.867	0.867	0.867
1.5		0.886	0.875	0.903	0.469	0.472	0.353	1.5		0.861	0.847	0.850	0.850	0.850	0.850
2.0		0.733	0.919	0.919	0.925	0.925	0.922	2.0		0.844	0.842	0.844	0.844	0.844	0.844
2.5	1	0.925	0.925	0.931	0.919	0.922	0.922	2.5	5	0.847	0.836	0.844	0.842	0.836	0.833
3.0		0.925	0.925	0.922	0.919	0.919	0.919	3.0		0.861	0.853	0.856	0.858	0.853	0.856
3.5		0.919	0.919	0.917	0.917	0.917	0.919	3.5		0.858	0.856	0.853	0.858	0.853	0.853
4.0		0.914	0.917	0.917	0.917	0.917	0.919	4.0		0.856	0.856	0.853	0.858	0.853	0.850

23

Time frame/(s)	K^{in}	Interval time/(s)						Time frame/(s)	K^{in}	Interval time/(s)					
		1	2	3	4	5	6			1	2	3	4	5	6
4.5		0.914	0.917	0.917	0.914	0.917	0.922	4.5		0.856	0.847	0.853	0.850	0.850	0.850
5.0		0.914	0.917	0.917	0.914	0.919	0.922	5.0		0.853	0.853	0.853	0.850	0.850	0.850
1.0		0.669	0.658	0.453	0.850	0.856	0.897	1.0		0.119	0.083	0.086	0.122	0.128	0.103
1.5		0.928	0.925	0.928	0.925	0.928	0.925	1.5		0.842	0.856	0.861	0.858	0.858	0.861
2.0		0.939	0.931	0.928	0.928	0.928	0.928	2.0		0.817	0.828	0.825	0.819	0.825	0.825
2.5		0.933	0.931	0.931	0.933	0.931	0.931	2.5		0.800	0.794	0.783	0.789	0.789	0.792
3.0	2	0.931	0.936	0.939	0.936	0.931	0.931	3.0	6	0.797	0.794	0.797	0.789	0.789	0.792
3.5		**0.933**	**0.939**	**0.933**	**0.933**	**0.936**	**0.936**	3.5		0.794	0.789	0.794	0.794	0.797	0.797
4.0		**0.936**	**0.942**	**0.939**	**0.936**	**0.933**	**0.936**	4.0		0.794	0.794	0.797	0.800	0.800	0.800
4.5		**0.925**	**0.936**	**0.933**	**0.936**	**0.936**	**0.936**	4.5		0.800	0.803	0.803	0.803	0.800	0.803
5.0		0.928	0.928	0.933	0.933	0.933	0.936	5.0		0.797	0.797	0.794	0.792	0.789	0.789
1.0		0.914	0.914	0.911	0.911	0.908	0.903	1.0		0.819	0.814	0.814	0.819	0.822	0.817
1.5		0.908	0.906	0.906	0.900	0.900	0.900	1.5		0.839	0.847	0.842	0.844	0.850	0.847
2.0		0.892	0.889	0.886	0.886	0.886	0.886	2.0		0.794	0.789	0.789	0.794	0.789	0.794
2.5		0.881	0.878	0.878	0.878	0.869	0.869	2.5		0.803	0.794	0.783	0.775	0.769	0.767
3.0	3	0.883	0.886	0.886	0.886	0.886	0.883	3.0	7	0.778	0.778	0.775	0.772	0.761	0.761
3.5		0.892	0.881	0.878	0.878	0.878	0.878	3.5		0.767	0.769	0.767	0.764	0.764	0.761
4.0		0.892	0.883	0.883	0.878	0.878	0.878	4.0		0.767	0.764	0.761	0.764	0.761	0.761
4.5		0.889	0.886	0.881	0.881	0.881	0.889	4.5		0.767	0.767	0.758	0.761	0.758	0.756
5.0		0.886	0.886	0.886	0.889	0.886	0.883	5.0		0.764	0.758	0.761	0.756	0.764	0.756
1.0		0.897	0.900	0.903	0.906	0.903	0.897								
1.5		0.878	0.886	0.892	0.892	0.892	0.889								
2.0		0.875	0.881	0.881	0.881	0.881	0.881								
2.5		0.881	0.875	0.878	0.872	0.872	0.869								
3.0	4	0.869	0.869	0.872	0.867	0.864	0.861								
3.5		0.867	0.864	0.864	0.867	0.864	0.861								
4.0		0.858	0.856	0.861	0.864	0.858	0.853								
4.5		0.856	0.858	0.858	0.856	0.856	0.858								
5.0		0.858	0.856	0.856	0.856	0.856	0.856								

Table S3 Summary of accuracy under different training parameters

References:

1. X. F. Li, L. Basile, B. Huang, C. Ma, J. W. Lee, I. V. Vlassiouk, A. A. Puretzky, M. W. Lin, M. Yoon, M. F. Chi, J. C. Idrobo, C. M. Rouleau, B. G. Sumpter, D. B. Geohegan and K. Xiao, *ACS Nano*, 2015, **9**, 8078-8088.
2. I. Susanto, C.-Y. Tsai, Y.-T. Ho, P.-Y. Tsai and I.-S. Yu, *Nanomaterials*, 2021, **11**, 1406.

3. Y. L. Zheng, W. L. Wang, X. C. Li, Y. Li, L. G. Huang and G. Q. Li, *Phys. Chem. Chem. Phys.*, 2017, **19**, 21467-21473.

4. Y. F. Han, Y. B. Dai, J. Wang, D. Shu and B. D. Sun, *Appl. Surf. Sci.*, 2011, **257**, 7831-7836.

5. C. L. Hu, Y. Chen, J. Q. Li and Y. F. Zhang, *Appl. Surf. Sci.*, 2008, **254**, 6514-6520.

6. E. N. Voloshina, D. Mollenhauer, L. Chiappisi and B. Paulus, *Chem. Phys. Lett.*, 2011, **510**, 220-223.

7. A. Soon, L. Wong, M. Lee, M. Todorova, B. Delley and C. Stampfl, *Surf. Sci.*, 2007, **601**, 4775-4785.

8. Y. Lu, J. Chen, T. Chen, Y. Shu, R. J. Chang, Y. Sheng, V. Shautsova, N. Mkhize, P. Holdway, H. Bhaskaran and J. H. Warner, *Adv. Mater.*, 2020, **32**, 1906958.

9. J. L. Tang, X. Yang, Y. Y. Wang, H. Wang, Y. Xiao, M. Apreutesei, Z. Nie and B. Normand, *Metals*, 2019, **9**, 294.

10. J. T. Yates, Jr, *Phys. Today*, 2004, **57**, 79-80.

11. Z. Lin, T. Lin, T. Lin, X. Tang, G. Chen, J. Xiao, H. Wang, W. Wang and G. Li, *Appl. Phys. Lett.*, 2023, **122**.

12. J. P. Zeng, Y. Zhang, Y. H. Chen, Z. J. Han, X. M. Chen, Y. Peng, L. Chen and S. Chen, *J. Mol. Graph. Model.*, 2022, **114**, 108184.

13. Q. Li, Y. Xiao, X. Shi and S. Song, *Nanomaterials*, 2017, **7**, 265.

14. Q. S. Bai, Y. H. Dou, X. He, A. M. Zhang and Y. B. Guo, *Acta Phys. SIN-CH ED.*, 2020, **69**, 226102.

15. Z. Taheri and A. N. Pour, *J. Mol. Model.*, 2021, **27**, 59.

16. G. N. Parsons and R. D. Clark, *Chem. Mater.*, 2020, **32**, 4920-4953.

17. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.

18. E. Pop, V. Varshney and A. K. Roy, *Mrs Bull.*, 2012, **37**, 1273-1281.

19. E. Ziade, J. Yang, G. Brummer, D. Nothern, T. Moustakas and A. J. Schmidt, *Appl. Phys. Lett.*, 2017, **110**, 031903.

20. Y. L. Sun, L. Qian, D. Xie, Y. X. Lin, M. X. Sun, W. W. Li, L. M. Ding, T. L. Ren and T. Palacios, *Adv. Funct. Mater.*, 2019, **29**, 1902538.

21. W. X. Wang, S. Gao, Y. Li, W. J. Yue, H. Kan, C. W. Zhang, Z. Lou, L. L. Wang and G. Z. Shen, *Adv. Funct. Mater.*, 2021, **31**, 2101201.

22. Z. E. Xia, J. R. Wang and Z. E. Fei, *Phys. Status Solidi-R*, 2019, **13**, 1900082.

23. L. Yang, M. Singh, S. W. Shen, K. Y. Chih, S. W. Liu, C. I. Wu, C. W. Chu and H. W. Lin, *Adv. Funct. Mater.*, 2021, **31**, 2008259.

24. Y. Yan, J. H. Yang, J. Du, X. M. Zhang, Y. Y. Liu, C. X. Xia and Z. M. Wei, *Adv. Mater.*, 2021, **33**, 2008761.

25. X. L. Zhang, J. Li, Z. Y. Ma, J. Zhang, B. Leng and B. D. Liu, *ACS Appl. Mater. Inter.*, 2020, **12**, 47721-47728.

26. Y. Lu, T. X. Chen, N. Mkhize, R. J. Chang, Y. W. Sheng, P. Holdway, H. Bhaskaran and J. H. Warner, *ACS Nano*, 2021, **15**, 19570-19580.

27. J. Bai, W. W. Xie, W. Q. Zhang, Z. P. Yin, S. S. Wei, D. H. Qu, Y. Li, F. W. Qin, D. Y. Zhou and D. J. Wang, *Appl. Surf. Sci.*, 2022, **600**, 154084.

28. J. Y. Li, Y. Z. Qian, W. Li, S. C. Yu, Y. X. Ke, H. W. Qian, Y. H. Lin, C. H. Hou, J. J. Shyue, J. Zhou, Y. Chen, J. P. Xu, J. T. Zhu, M. D. Yi and W. Huang, *Adv. Mater.*, 2023, **35**, 2209728.
29. F. Gul, *Ceram. Int.*, 2018, **44**, 11417-11423.
30. F. C. Chiu, *Adv. Mater. Sci. Eng.*, 2014, **2014**, 578168.
31. X. Qin, H. Liu, J. Hu, J. Huang, F. Yang, B. Sun, Y. Zhao, M. Xu, X. Duan, M. Huang and Y. Zhang, *J. Electron. Mater.*, 2023, **52**, 3868-3876.
32. Y. B. Deng, X. G. Xu, Z. D. Xu, M. X. Wang, Q. Liu, Y. L. Ma, J. K. Chen, K. K. Meng, Y. Wu, J. Miao and Y. Jiang, *Ceram. Int.*, 2022, **48**, 4693-4698.
33. P. Steiner, A. Jalalvand, S. Stone and P. Birkholz, *Eng. Appl. Artif. Intell.*, 2022, **113**, 104964.