Advanced Optical Materials

ADVANCED OPTICAL MATERIALS
www.advopticalmat.de

RESEARCH ARTICLE

A Biomimetic Binocular Visual System Based on β-Ga_2O_3 Optoelectronic Synaptic Device for Motion Recognition

Haolan Qu[1,2,3] | Haodong Jiang[1,2,3] | Changshu Liu[1] | Kerui Li[4] | Hongzhi Wang[4] | Tiangui You[2,3] | Xin Ou[2,3] | Wenhui Xu[2,3] | Xinbo Zou[1]

[1]School of Information Science and Technology, ShanghaiTech University, Shanghai, People's Republic of China | [2]Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, People's Republic of China | [3]School of Microelectronics, University of Chinese Academy of Sciences, Beijing, People's Republic of China | [4]College of Materials Science and Engineering, Donghua University, Shanghai, People's Republic of China

Correspondence: Wenhui Xu (xuwh@mail.sim.ac.cn) | Xinbo Zou (zouxb@shanghaitech.edu.cn)

Received: 28 April 2026 | **Revised:** 24 July 2026 | **Accepted:** 5 August 2026

Keywords: β-Ga_2O_3 | biomimetic binocular visual system | motion recognition | optoelectronic artificial synapse

ABSTRACT

With the rapid development of autonomous vehicles, robotics, and monitoring systems, the demand for having visual system capable of recognizing explicit motion information, such as motion direction, state, and speed, has increased substantially. To address this demand, biomimetic visual systems have garnered considerable attention owing to their high speed and capability of collecting intricate motion details. However, most biomimetic visual systems are limited to two-dimensional (2D) recognition via monocular manner. In contrast, binocular visual systems could extend the recognition scope from the 2D plane to three-dimensional (3D) space. This study proposes a biomimetic binocular visual system for 3D motion recognition based on β-Ga_2O_3 synaptic devices to gather motion information including motion direction, state, and speed. Based on trap behaviors, the broad tunability of synaptic behavior endows β-Ga_2O_3 synaptic device with multiple functions including photodetector (PD), compressor and activation function generator (AFG), realizing the recognition of nineteen directions in 3D space with a high accuracy of 98.83%. Furthermore, two additional applications, namely motion state and speed recognition, are also validated by this system. These results pioneer the development of a biomimetic binocular visual system and paves a solid path for future 3D motion recognition.

1 | Introduction

In recent years, the rapid development of autonomous vehicles, robotics, and monitoring systems has created a growing demand for visual systems capable of detecting moving objects [1–3]. Biomimetic visual systems have garnered considerable research interest because of their high-speed performance and ability to capture detailed motion information, such as motion direction, state, and speed. Unlike static object recognition, moving-object identification generally involves the acquisition and processing of temporal information. Previously, a typical approach was to acquire multiple responses at different time points using photodetector (PD) arrays and combine them to construct a time-series dataset. For example, temporal information was generated by superimposing two consecutive responses acquired with an 18 × 18 ripple-assisted optoelectronic array, thereby enabling all-day motion detection [4]. Nevertheless, massive amounts of data would be generated, leading to challenge in data storage and computation. Recently, multiple consecutive signal bits have been compressed into one single output response using a biomimetic visual system inspired by flying insects [2, 5, 6]. Via this biomimetic visual system, explicit motion information, such as motion direction and speed, can be effectively extracted with only one single sequence of output bits [2].

Most existing visual systems for motion recognition typically operate in a two-dimensional (2D) monocular manner, capturing 2D projection of moving objects, thereby restricting the acquired motion information within a 2D plane [5–9]. Although four fundamental directions along horizontal and vertical axes in a 2D plane have been successfully analyzed and identified, directions beyond the particular plane remain unexplored [5]. Growing research attention has been directed toward expanding the recognition scope from a 2D plane to three-dimension (3D) space. Inspired by flying insects, Chen et al. developed a biomimetic visual system capable of generating compressed spatiotemporal representations of motion in seven directions distributed across three mutually perpendicular axes in 3D space, demonstrating its potential for 3D motion sensing [2]. However, in the subsequent recognition experiments, the artificial neural network (ANN) model was used to identify only five motion directions confined to a 2D plane, while out-of-plane motion along the third axis was not included. Therefore, the system's capability for full 3D motion recognition remained unverified, potentially raising safety concerns in real-world applications such as autonomous driving. Although motional objects in 3D space have been successfully identified via high-contrast bidirectional optoelectronic synaptic devices, the detailed motion attributes including motion direction and speed were still missing [10]. The limited amount of information provided by monocular 2D imaging constitutes a major obstacle to comprehensive 3D motion recognition. In contrast, a biomimetic binocular visual system can acquire complementary information from two distinct 2D projections, offering a promising approach to extending motion recognition from a 2D plane to 3D space and enabling the determination of additional motion attributes, including 3D direction, state, and speed.

Toward the goal of implementing an intelligent biomimetic visual system, a variety of device technologies have been developed to fulfill the demands of key building components in visual systems, including PD, compressor, differentiator, and so on [2, 5, 10]. Based on high photoresponsivity, Ga_2O_3/In_2Se_3-based sensors have been employed as PD in motion direction recognition, successfully enabling the identification of four directions [6]. In the preprocessing part, both compressor and differentiator have been successfully incorporated using optically controlled bidirectional synaptic behavior of PbS/C60/Pentacene device [5]. Memory function (ME) has been achieved by leveraging self-assembled P3HT structure with Schottky barriers to form exceeding 128 memory states within the solar-blind range [11]. Weight mapping (WM) has also been implemented for motion detection, using $BP/Al_2O_3/WSe_2$/h-BN heterostructure, benefiting from its excellent linearity of the multilevel conductance state [7]. Recently, trap-state-mediated synaptic devices have attracted significant research interest due to their low power consumption, high repeatability, high regulatory flexibility of synaptic behaviors, and sensitive response to optical pulse, highlighting their considerable potential for implementation in visual system. Furthermore, synaptic device can function as compressor due to its characteristic responses to multiple optical pulses [5]. Activation function generator (AFG), providing indispensable nonlinear computing parts, also plays a crucial role in neutral network. Based on the nonlinear relationship between input voltage bias and excitatory postsynaptic current (EPSC), trap-state-mediated synaptic devices are capable of mimicking a tailor-made activation function, offering a potential alternative to the rectified linear unit (ReLU) function and accelerating the training process [12].

In this paper, a biomimetic binocular visual motion recognition system based on β-Ga_2O_3 synaptic devices is developed. Under optical pulses, synaptic features such as EPSC, paired-pulse facilitation (PPF) and voltage-modulated current amplitude are demonstrated. For moving objects, nineteen directions in 3D space are successfully identified based on the biomimetic binocular visual system. With the synaptic behaviors, β-Ga_2O_3 synaptic devices function as a photodetector, compressor and activation function generator not only in the preprocessing part but also in the neural network. Furthermore, two other functions, motion state recognition and motion speed recognition, are validated by this biomimetic binocular visual system as well. These results provide a promising foundation for the future development of biomimetic binocular visual systems.

2 | Results and Discussion

The manufactory process began with the wafer fabrication, which contained ion implantation, wafer bonding, annealing, exfoliation and polishing. A 2-inch β-Ga_2O_3 wafer with a shallow damage layer via hydrogen ions implantation was bonded with a 4-inch 4H-SiC wafer. After annealing at 450°C, a β-Ga_2O_3 thin film was exfoliated to the 4H-SiC wafer. The β-Ga_2O_3/SiC heterogeneous integrated wafer was finally achieved after a modified chemical-mechanical polishing (CMP) process and the high-temperature treatment. Mesa isolation was formed by inductively coupled plasma-reactive ion etching (ICP-RIE) and silicon ions were implanted into β-Ga_2O_3 conductive channel to improve the carrier concentration. The enhancement mode β-Ga_2O_3 metal-oxide-semiconductor field effect transistor (MOSFET) was realized by gate recess process. Electron-beam evaporation (EBE) and rapid thermal annealing (RTA) were applied to form ohmic contact of Ti/Au (20/160 nm). The gate stack contained a 30-nm-thick Al_2O_3 layer and a Ni/Au (20/160 nm) Schottky contact deposited by atomic layer deposition (ALD) and EBE, respectively. More detailed information on the fabrication of devices was published in previous work [13], and the fabrication process together with the image of β-Ga_2O_3 MOSFET are plotted in supporting information (Figures S1 and S2).

The fabricated β-Ga_2O_3 MOSFET features gate length (L_G)/gate-source distance (L_{GS})/gate-drain distance (L_{GD})/gate width (W_G) of 4/4/20/20 µm. Using the linear extrapolation method, the threshold voltage (V_{th}) extracted from the transfer curves is determined to be 15.53 V at a drain voltage (V_{DS}) of 9 V. An extremely low gate leakage current (I_G) level smaller than 10 pA validates the excellent reverse blocking characteristics of the β-Ga_2O_3 MOSFET. The optical image and detailed electrical characteristics of β-Ga_2O_3 MOSFET are shown in supporting information (Figure S3).

As shown in Figure 1a, in biological systems, synapses act as the channel to transmit biological information between two neurons, and the transmission start with the exocytosis of neurotransmitters from presynaptic neurons to synaptic clefts upon the biological stimulus, followed by the acceptance of

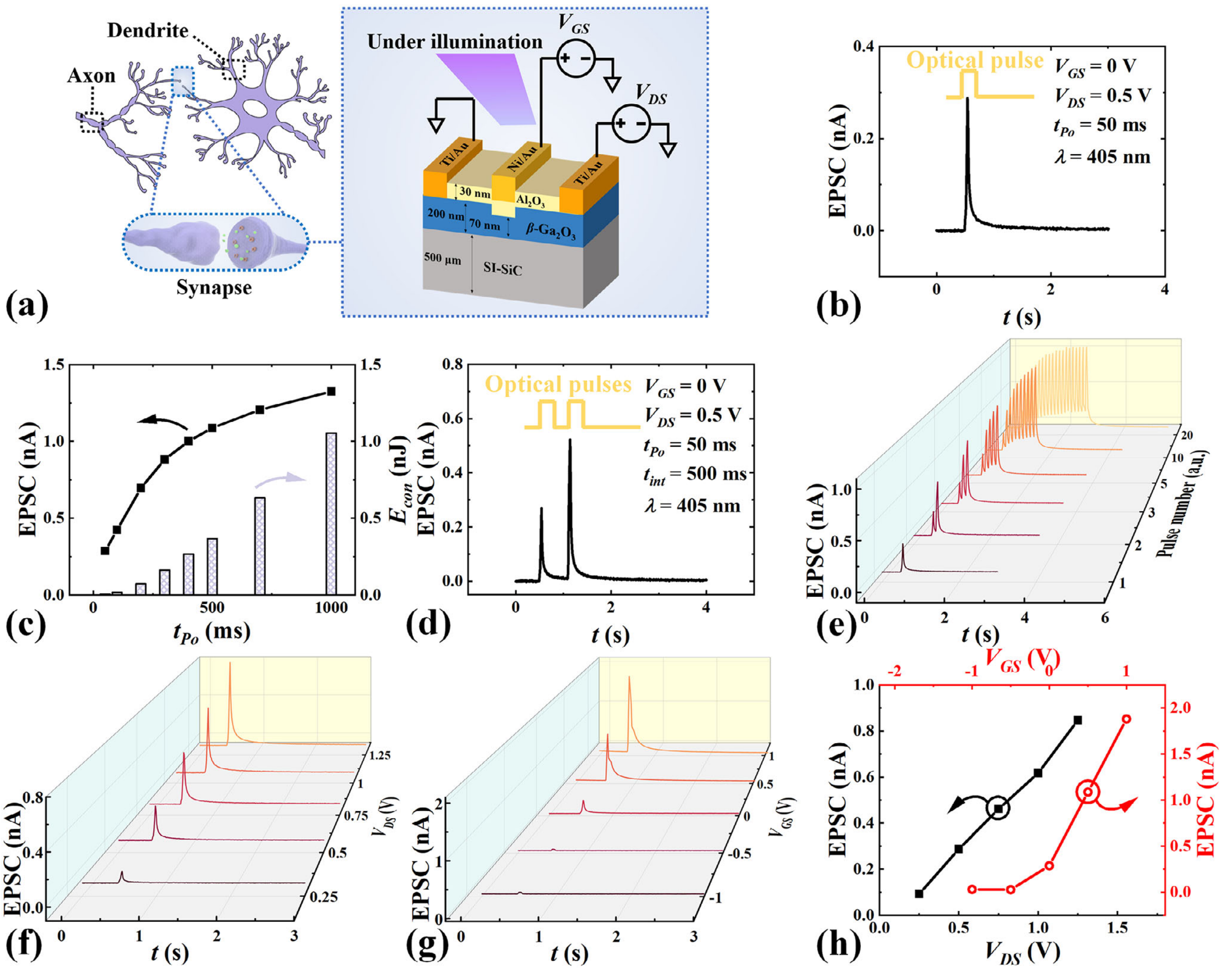

FIGURE 1 | (a) The schematic of a biological synapse and β-Ga_2O_3 synaptic device. (b) EPSC stimulated by an optical pulse. (c) EPSC and energy consumption (E_{con}) with various optical pulse widths (t_{Po}). (d) EPSC stimulated by two adjacent optical pulses. (e) EPSC stimulated by optical pulses with pulse numbers from 1 to 20. (f) EPSC at gate voltage (V_{GS}) of 0 V. (g) EPSC at V_{DS} of 0.5 V. (h) Peak EPSC value modulated by V_{DS} and V_{GS}.

neurotransmitters by receptors on postsynaptic neurons, resulting in EPSC [14–19]. In order to emulate the biological synapse, the gate electrode of the β-Ga_2O_3 MOSFET acts as the presynaptic neuron, and the β-Ga_2O_3 conductive channel together with drain electrode and source electrode of the β-Ga_2O_3 MOSFET operates as the postsynaptic neuron.

As illustrated in Figure 1b, EPSC of β-Ga_2O_3 synaptic device is triggered by an optical pulse spiking with a t_{Po} of 50 ms and a wavelength (λ) of 405 nm, together with V_{GS} and V_{DS} of 0 and 0.5 V, respectively, achieving a peak current amplitude of 287 pA upon the optical pulse stimulation. During the optical pulse, photogenerated electron–hole pairs are either captured by trap states or accumulated in β-Ga_2O_3 conductive channel, and the resulting photoexcited current, donated as EPSC, is formed by the accumulated carriers and positive V_{DS}. Compared to the sharp increase, the current gradually reduces subsequent to the excitation due to the relatively slow release process of carriers from trap states. The high repeatability of EPSC with errors below 10% and photoexcited current with various t_{Po} are shown in supporting information (Figure S4). As exhibited in Figure 1c, EPSC increases from 287 to 1325 pA with t_{Po} rising from 50 to 1000 ms, proving an enhanced accumulation in β-Ga_2O_3 conductive channel with increasing t_{Po}, and the corresponding E_{con} of EPSC can be calculated by V_{DS} and drain current (I_{DS}): [20, 21]

$$\int dE_{con} = \int V_{DS} \times I_{DS} \times dt \tag{1}$$

where, t represents the time during the optical pulse. With an optical pulse spanning for 50 ms, an extremely low E_{con} of 4.67 pJ is derived, which raises to 1054.23 pJ when t_{Po} extends to 1000 ms. Additionally, on/off ratio, signal-to-noise ratio (SNR) and the comparison of EPSC of synaptic devices are displayed in supporting information (Figure S5 and Table S1).

PPF, a typical temporal enhancement in synaptic connection, reflects the ability to process the temporal signals in biology system, and two continuous optical pulses are regarded as stimuli to mimic biological PPF [22–24]. As exhibited in Figure 1d, PPF index, defined as the ratio of peak amplitude of second EPSC (A_2) to that of first EPSC (A_1), reaches 1.93 with a t_{Po} of 50 ms and an interval time (t_{int}) of 500 ms, which can be attributed to trap behavior. When the second optical pulse is applied before the trapped carriers generated by the first optical pulse are fully released, fewer empty trap states remain available to capture carriers during the second optical pulse, and a heightened number of carriers are accumulated in the β-Ga_2O_3 conductive channel, leading to an enhanced A_2. The detailed fitting results of PPF match well with the biological synapse and prove an excellent short-term memory (STM), as shown in supporting information (Figure S6) [12, 25].

With continuous rehearsal of optical pulses, the memory effect of the β-Ga_2O_3 synaptic device is consolidated, leading to a transition from short-term memory (STM) to long-term memory (LTM) [26, 27]. Figure 1e shows EPSC induced by optical pulses with pulse numbers from 1 to 20, and EPSC rises due to the enhanced accumulation of photoexcited carriers in the β-Ga_2O_3 conductive channel with increasing optical pulse number. When the optical pulse number is smaller than 5, photogenerated carriers are primarily captured by shallow trap states in the β-Ga_2O_3 MOSFET, thus forming STM, while a larger fraction of photogenerated carriers are captured by deep level trap states which require more time to release carriers in the β-Ga_2O_3 MOSFET, thereby establishing LTM when the optical pulse number extends to 20. The schematic of LTM and peak EPSC value as a function of optical pulse number is shown in supporting information (Figure S7). Meanwhile, the learning-forgetting-relearning effect demonstrating an outstanding memory effect is shown in supporting information (Figure S8).

Figure 1f,g illustrates EPSC stimulated by optical pulses with various V_{DS} and V_{GS}, respectively, and the corresponding peak amplitude value of EPSC is plotted in Figure 1h. When V_{DS} spans from 0.25 to 1.25 V in increments of 0.25 V, the peak EPSC amplitude increases from 92 to 847 pA with a significant enhancement of 820.65%, indicating the improved generation process of electron–hole pairs due to the larger electric field from drain to source. More results on EPSC with V_{DS} of 0.25 V with t_{Po} from 50 to 1000 ms are shown in supporting information (Figure S9). When V_{GS} is positive, the peak EPSC amplitude increases with the rising V_{GS}, while it is significantly suppressed with negative V_{GS}, and the dependence of EPSC on V_{GS} is governed by trap behaviors. With a positive V_{GS}, narrow depletion region together with the positive electric field from gate to source leads to the filled trap states, and only a number of traps states capture carriers during optical pulses, contributing to the accumulation in the β-Ga_2O_3 conductive channel. With a rising positive V_{GS}, a larger gate-source electric field strength is formed, leading to an augment of generation of electron−hole pairs and improved accumulation of carriers, thus increasing EPSC. On the contrary, a negative V_{GS} creates a negative electric field from gate to source and widens the depletion region, resulting in the empty trap states which capture the photogenerated carriers before they are accumulated in β-Ga_2O_3 MOSFET, thereby suppressing the EPSC. The similar trap behaviors with gate voltage stress are also observed in previous studies [28, 29]. More results on EPSC with V_{GS} of −0.5 V and t_{Po} from 50 to 1000 ms are shown in supporting information (Figure S10).

Compared to a monocular visual recognition system, two projections of a 3D object are observed in a biological binocular visual recognition system, providing more information and resulting in a more powerful capability for recognition. Inspired by this biological mechanism, a biomimetic binocular visual recognition system based on β-Ga_2O_3 synaptic devices is proposed in this work, which is depicted in Figure 2a. In order to recognize more directions in 3D space, two distinct projections of the motional object are collected (the blue and green projections correspond to the views perceived by the left and right eyes, respectively). According to the coordinate system exhibited in Figure 2a, for example, if the object moves along $(0, y^+, 0)$, the blue projections detected along x^- direction on the y-z coordinate plane are observed to move toward the right, while the green projections detected along y^- direction on the x-z coordinate plane is still regardless the size of the object. Furthermore, the β-Ga_2O_3 synaptic devices are employed to compress three frames into one single frame preserving all temporal information from the three frames by encoding three continuous bits of the same pixel at three points in time as a three-bit optical pulse train. For instance, if the three-bit pixel exhibits [0 0 0], no optical pulse will be applied on the β-Ga_2O_3 synaptic device, while [1 1 1] represents three adjacent optical pulses. According to the normalization of the output current of the β-Ga_2O_3 synaptic device, the corresponding normalized gray-level is achieved, and the final input gray-level of the pixel is derived by inverting the corresponding normalized gray-level. By combining two compressed projections, the final image carrying temporal information of three frames and two projections is acquired, and Figure 2b shows four representative directions, along (x^+, y^-, z^+), $(x^-, 0, 0)$, $(0, y^-, 0)$, and (x^-, y^-, z^-), for recognition. Figure 2c displays some motion directions described by the coordinate system, in which the six simple directions along x^+, x^-, y^+, y^-, z^+ and z^- are hidden for clarity.

Figure 3a illustrates the schematic of the transformation process from optical pulse train to final input gray-level. Figure 3b,c shows the current responses to optical pulse train of [1 0 0] and [0 1 1] with a t_{Po} of 50 ms and a t_{int} of 50 ms, respectively, in which the current level is measured at the moment 250 ms after the onset of the first bit, yielding current level of 21 and 581 pA with optical pulse train of [1 0 0] and [0 1 1], respectively. Figure 3d illustrates the current responses of the β-Ga_2O_3 synaptic device corresponding to all three-bit optical pulse sequences. After each transformation process of the three-bit optical pulse sequence, the synaptic states of the devices would need to be initialized or reset for the next three-bit optical pulse sequence. For comparison, corresponding training images and testing images for the conventional PD and β-Ga_2O_3 synaptic device are both created. Different from the β-Ga_2O_3 synaptic device, conventional PD has no capability to process temporal information, and the corresponding images only consist of the information of the final frame. More detailed information of training images and testing images is shown in supporting information (Figure S11). As plotted in Figure 3e, the CNN for motion direction recognition consists of two convolution layers, two max-pool layers, and three full-connection layers, and the recognition accuracy with epochs from 1 to 30 is shown in Figure 3f. Only retaining the information of the final frame, the motion direction cannot be recognized absolutely based on conventional PD without ability to process temporal information with the accuracy maintaining around 20% and reaching 19.08% with an epoch of 30, while the accuracy based on the β-Ga_2O_3 synaptic device rapidly increases above 90% when epoch reaches 7, and achieves 97.86% with epoch of 30. Figure 3g,h exhibits the

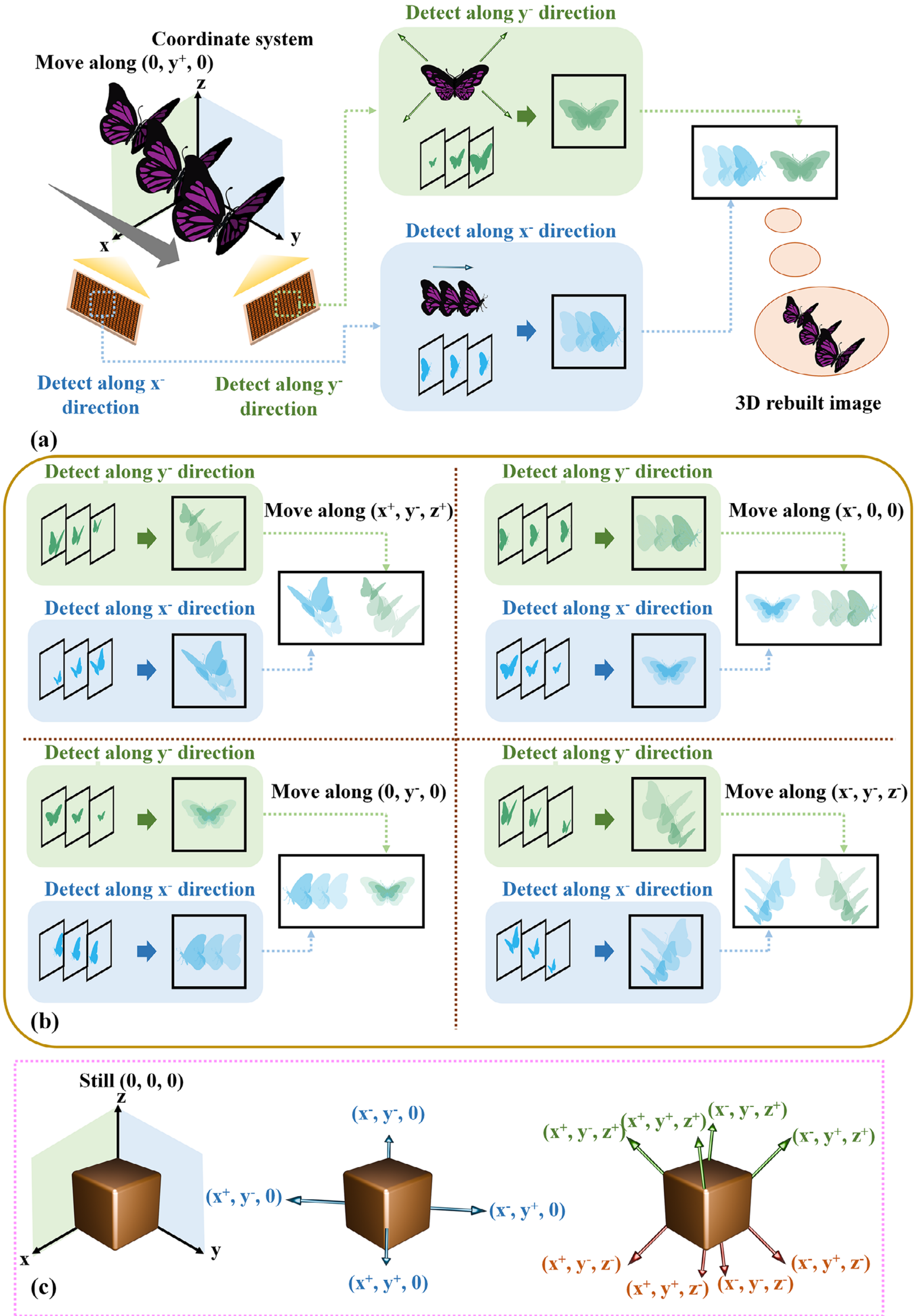

FIGURE 2 | (a) Biomimetic binocular visual motion recognition system based on β-Ga_2O_3 synaptic devices. (b) The schematic of four example directions for recognition. (c) The schematic of some directions for recognition.

confusion matrices of motion direction recognition with an epoch of 30 based on conventional PD and the β-Ga_2O_3 synaptic device, respectively, which further confirm the recognition results. For simplicity, the directions in the following confusion matrix are abbreviated, for example, (–, +, 0) represents along (x^-, y^+, 0). The confusion matrix based on conventional PD without ability to process temporal information contains almost no red grids along the diagonal, indicating the wrong recognition for nineteen motion direction, on the contrary, the confusion matrix based on the β-Ga_2O_3 synaptic device exhibits prominent red

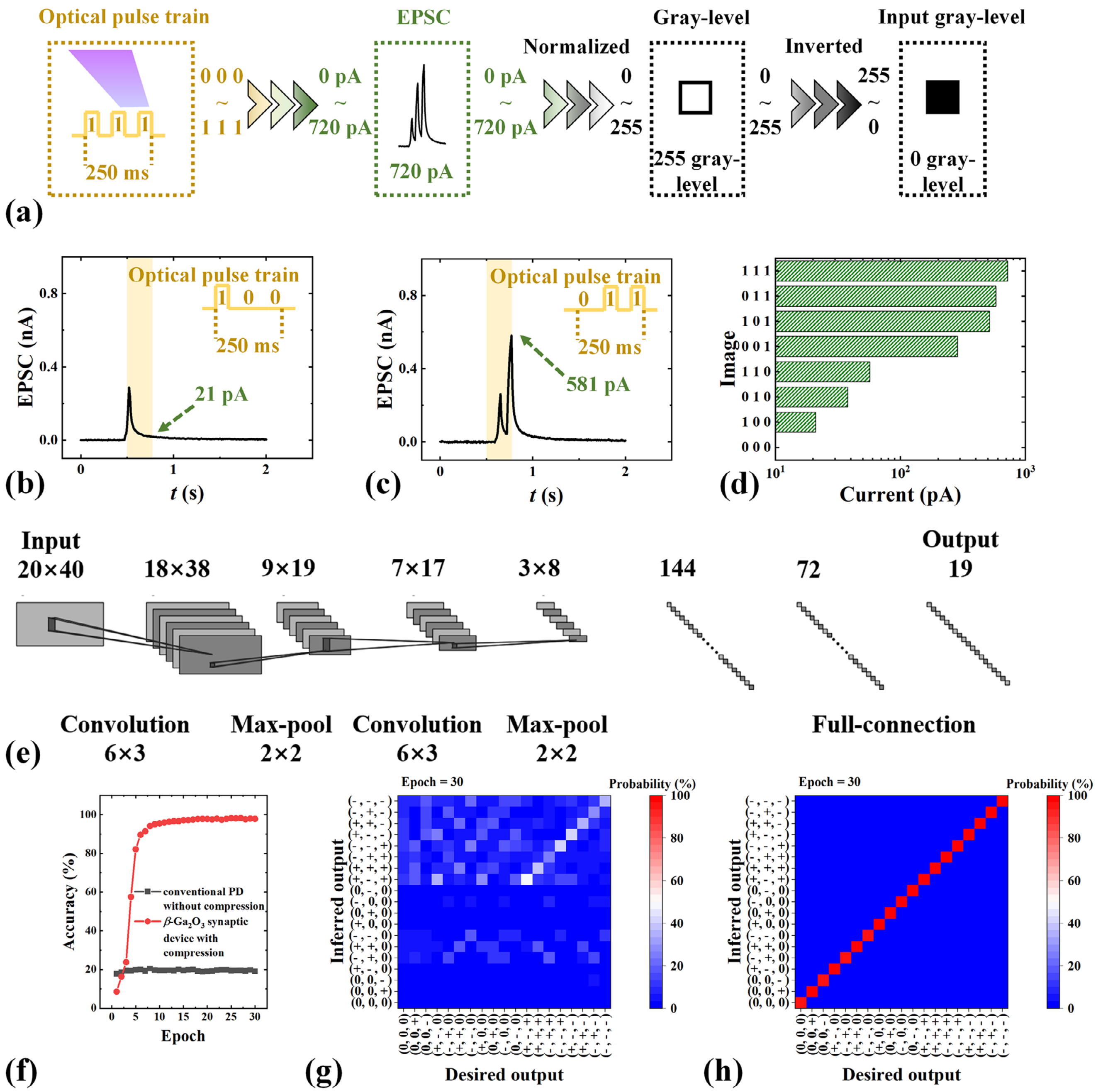

FIGURE 3 | (a) The schematic of the variation process from an optical pulse train to input gray-level. The optical pulse train of example (b) [1 0 0] and (c) [0 1 1]. (d) Current response to three-bit optical pulses. (e) The structure of a convolutional neural network (CNN). (f) The accuracy of the CNN based on conventional PD without the ability to process temporal information and β-Ga_2O_3 synaptic device. The confusion matrices based on (g) conventional PD without ability to process temporal information and (h) β-Ga_2O_3 synaptic device.

grids on the diagonal line, representing the excellent recognition performance. More detailed information of confusion matrices based on conventional PD and β-Ga_2O_3 synaptic device is shown in supporting information (Figures S12 and S13).

In addition to the detection and compression in a biomimetic binocular visual motion recognition system, β-Ga_2O_3 synaptic device can also act as AFG in the CNN model, as shown in Figure 4a. As illustrated in Figure 4b, the activation function introduces the essential nonlinear part during the computation in neural networks, which brings the ability for fitting complex curves to neural networks. ReLU function is one of the most common activation functions and can be describe below:

$$f(n) = \begin{cases} 0 \text{ if } n < 0 \\ n \text{ if } n \geq 0 \end{cases} \quad (2)$$

In neural network, ReLU function offers the advantages of simple computing, avoiding vanishing gradients and sparsity of the model. However, the ReLU function still suffers from certain limitations during training of neural networks, including the unsmooth gradient at zero point which leads to a longer training

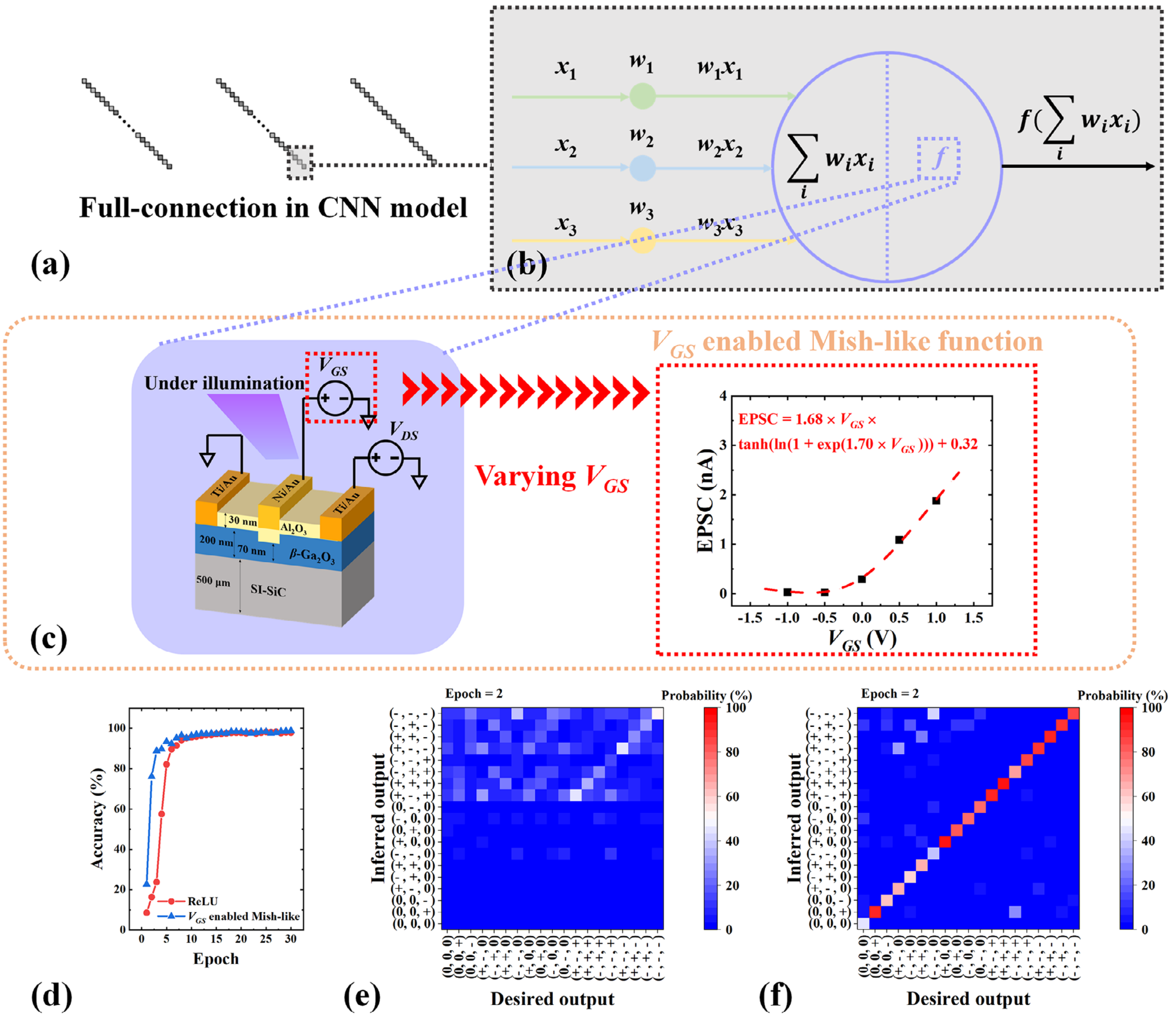

FIGURE 4 | (a) The schematic of full-connection in CNN model. (b) The schematic of single neuron in neural networks. (c) The activation function based on β-Ga_2O_3 synaptic devices. (d) The accuracy of the CNN based on the ReLU function and V_{GS} enabled Mish-like function. The confusion matrices based on (e) ReLU function and (f) V_{GS} enabled Mish-like function.

process and the dying neuron which contributes to the decreasing overall expressive ability of the model. As displayed in Figure 4c, a new V_{GS} enabled Mish-like function is obtained by modulating V_{GS}, and the fitting results of EPSC as a function of V_{GS} is derived below:

$$\mathrm{EPSC} = 1.68 \times V_{GS} \times \tanh(\ln(1 + \exp(1.70 \times V_{GS}))) + 0.32 \quad (3)$$

During the test, optical illumination with a t_{Po} of 50 ms is used to generate optical-induced carriers and form EPSC, together with V_{DS} of 0.5 V. The V_{GS} enabled Mish-like function possesses a smooth gradient, resulting in the improved training speed, together with the non-zero gradient of negative input, avoiding the dying neuron during training. For comparison, the same training images, testing images and CNN are employed to verify the effect of V_{GS} enabled Mish-like function, and the recognition accuracy with epochs from 1 to 30 is shown in Figure 4d. When the epoch extends to 30, the accuracy based on V_{GS} enabled Mish-like function reaches 98.83%, a little higher than that based on the ReLU function (97.86%). More importantly, the accuracy with an epoch of 2 using the ReLU function and V_{GS} enabled Mish-like function exhibits 16.36% and 76.03%, respectively, demonstrating that V_{GS} enabled Mish-like function exhibits a substantially faster training speed than the ReLU function due to the smooth gradient. Figure 4e,f displays the confusion matrices of motion direction recognition with an epoch of 2 based on the ReLU function and V_{GS} enabled Mish-like function, respectively. The confusion matrix with epoch of 2 using the ReLU function is dominated by blue and white grids, leading to a largely disordered classification, while several red grids exist along the diagonal based on V_{GS} enabled Mish-like function, indicating a much better recognition performance. More detailed information of confusion matrices based on V_{GS} enabled Mish-like function is shown in supporting information (Figure S14).

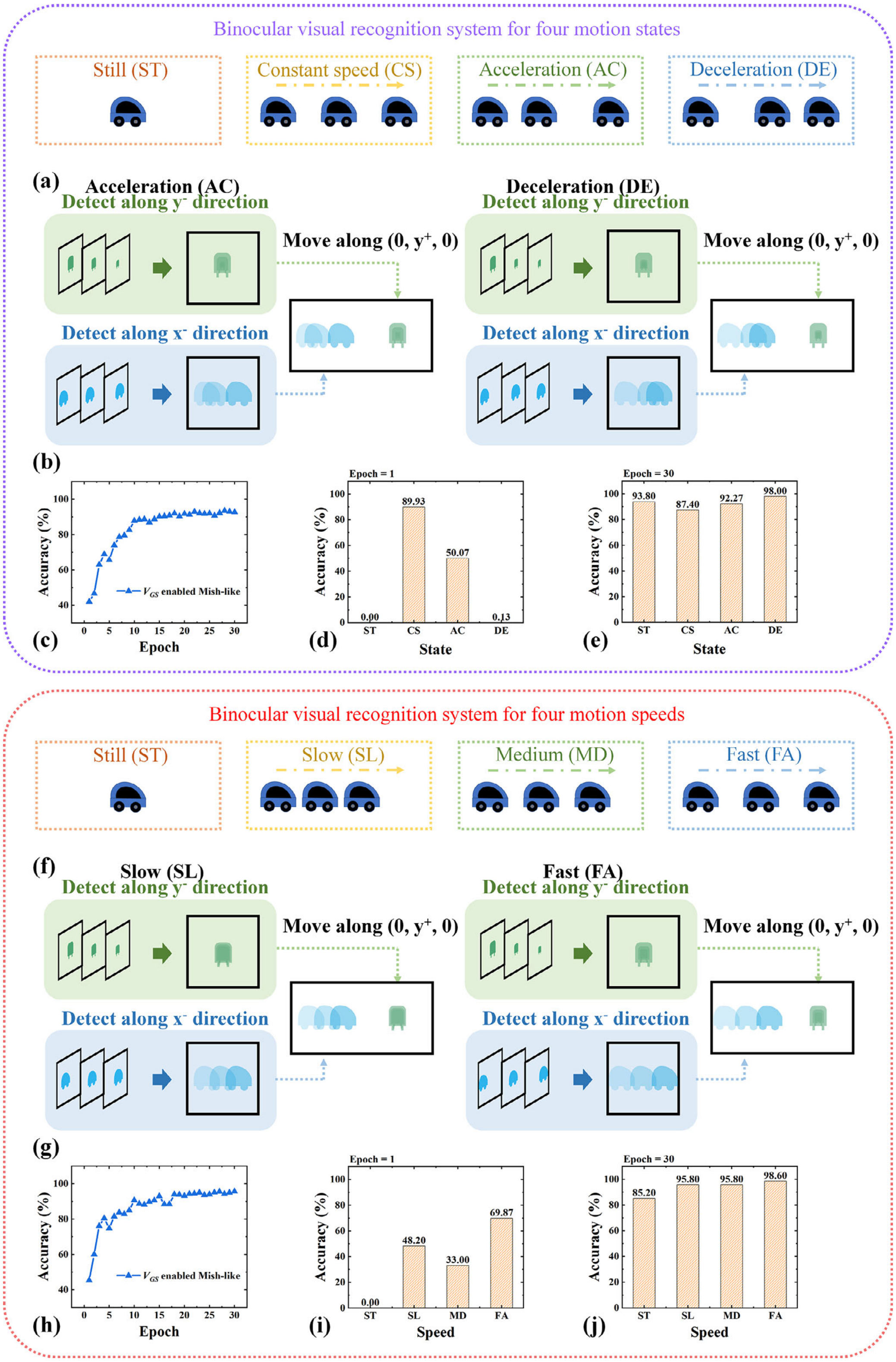

FIGURE 5 | (a) The schematic of four motion states for recognition. (b) The schematic of two example states for recognition. (c) The accuracy of the CNN based on β-Ga_2O_3 synaptic devices with V_{GS} enabled Mish-like function. The accuracy of each motion state with epochs of (d) 1 and (e) 30. (f) The schematic of four motion speeds for recognition. (g) The schematic of two example speeds for recognition. (h) The accuracy of the CNN based on β-Ga_2O_3 synaptic devices with V_{GS} enabled Mish-like function. The accuracy of each motion speed with epochs of (i) 1 and (j) 30.

TABLE 1 | The comparison of function of devices and application of visual systems for motion recognition.

Material	Dimentions	Function [a)]	Object	Direction	State	Speed
β-Ga_2O_3 (This work)	3D	CO and AFG		✓	✓	✓
MoS_2 [2]	2D	CO		✓		✓
PbS/C60/Pentacene [5]	2D	CO and DI		✓		
Ga_2O_3/ In_2Se_3 [6]	2D	CO		✓		
Perovskite [31]	2D	CO		✓	✓	✓
BP/Al_2O_3/WSe_2/h-BN [7]	2D	DI and WM	✓			
CdS/graphene/Ge [32]	2D	DI	✓			
MoS_2 [4]	2D	DI	✓			
P3HT/GaAs [11]	2D	ME	✓			
2D molecular crystal [10]	3D	CO	✓			

[a]To exhibit the table more concisely, the basic function of the photodetector (PD) realized by all reports is hidden. The abbreviation CO, AFG, DI, WM and ME represent compressor, activation function generator, differentiator, weight mapping, and memory effect, respectively.

Some previous reports have also investigated motion direction recognition systems and the comparison of motion direction recognition systems is shown in supporting information (Table S2) [2, 5, 6, 30]. Compared to the previous studies, the binocular visual motion recognition system based on β-Ga_2O_3 synaptic devices holds the capability to recognize nineteen directions in 3D space, together with high accuracy above 98%. Although the present work is confined to simulation of a visual system, the results substantiate the potential and feasibility of a biomimetic binocular architecture based on trap-state-mediated β-Ga_2O_3 devices. Realizing a hardware implementation represents a key priority for subsequent research.

In addition to motion direction, more motion information can also be recognized by the binocular visual recognition system, such as motion state and speed. Figure 5a illustrates the schematic of four motion states: still, uniform speed, acceleration, and deceleration, and Figure 5b plots the schematic of two example states for recognition. More detailed information of training images and testing images is shown in supporting information (Figure S15). As plotted in Figure 5c, the accuracy using β-Ga_2O_3 synaptic devices with V_{GS} enabled Mish-like function achieves 92.68%, demonstrating the ability of this system to handle complex mission. As plotted in Figure 5d,e, the accuracy of still and deceleration remain extra low, below 1% with an epoch of 1, while all motion states hold high accuracy above 85% with an epoch of 30, demonstrating the effective training process. More detailed information of confusion matrices for β-Ga_2O_3 synaptic devices with V_{GS} enabled Mish-like function is shown in supporting information (Figure S16). Figure 5f illustrates the schematic of four motion speeds, and Figure 5g plots the schematic of two example speeds for recognition. More detailed information of training images, testing images, and velocity estimation is shown in supporting information (Figure S17). As exhibited in Figure 5h, when the epoch extends to 30, the recognition accuracy based on the β-Ga_2O_3 synaptic device with V_{GS} enabled Mish-like function reaches 95.58%, proving the excellent performance of the binocular visual motion recognition system. As plotted in Figure 5i,j, extra low accuracy below 1% is also observed of still with epoch of 1, while all motion speeds hold high accuracy above 85% with an epoch of 30, validating the effective CNN model. More detailed information of derivation of the proposed V_{GS} enabled Mish-like function and confusion matrices for β-Ga_2O_3 synaptic devices with V_{GS} enabled Mish-like function is shown in supporting information (Figure S18).

Table 1 summarizes the function of devices and application of visual systems in previous reports. Most existing visual systems for motion recognition are limited to 2D recognition [2, 4–7, 11, 31, 32], and only a few studies focus on 3D moving objects, while the recognition of detailed motion information such as direction, state, and speed is still missing [10]. On the contrary, biomimetic binocular visual recognition system in this work recognizes the detailed motion information containing motion direction, state and speed in 3D space. Moreover, in addition to PD, most of devices operate as a compressor and differentiator in preprocessing part instead of in the neural network [2, 4–6, 10, 31, 32]. Some reports demonstrated that devices can work as WM and ME in a neural network, while the detailed motion information such as direction, state, and speed is unavailable [7, 11]. In contrast, a β-Ga_2O_3 synaptic device can work not only as a compressor, but also as AFG in a neural network, demonstrating the multifunction of the β-Ga_2O_3 synaptic device. Furthermore, a biomimetic binocular visual recognition system based on β-Ga_2O_3 synaptic device demonstrates three distinct applications for detailed motion information, indicating the huge potential for employment in biomimetic systems.

3 | Conclusion

In summary, a biomimetic binocular visual system for motion recognition based on β-Ga_2O_3 synaptic devices is proposed and demonstrated. The device exhibits outstanding synaptic behaviors including EPSC with low E_{con} and high noise immunity, PPF, transformation from STM to LTM, and voltage-modulated current amplitude. By leveraging these synaptic characteristics, the system successfully recognizes motion in nineteen directions in 3D space, achieving accuracies of 97.86% and 98.83% using the ReLU and V_{GS} enabled Mish-like functions, respectively.

Moreover, the system successfully recognizes two additional motion attributes—motion state and speed—with accuracies of 92.68% and 95.58%, respectively. These results highlight the considerable potential of the proposed binocular motion recognition system for future biomimetic vision applications.

Author Contributions

Kerui Li: Writing – review and editing, funding acquisition. **Xinbo Zou**: writing – review and editing, resources, supervision, funding acquisition, project administration, conceptualization. **Haodong Jiang**: writing – review and editing, investigation, methodology. **Changshu Liu**: software, visualization, writing – review and editing. **Hongzhi Wang**: writing – review and editing, funding acquisition. **Tiangui You**: funding acquisition, resources, writing – review and editing. **Haolan Qu**: writing – original draft, writing – review and editing, software, conceptualization, investigation, data curation, validation, visualization, methodology, formal analysis. **Wenhui Xu**: funding acquisition, resources, writing – review and editing. **Xin Ou**: funding acquisition, resources, writing – review and editing.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 52131303, 6240033389 and 62404236) and Science and Technology Commission of Shanghai Municipality (No. 24DP1500300).

Funding

The National Natural Science Foundation of China 52131303, the National Natural Science Foundation of China 6240033389, the National Natural Science Foundation of China 62404236, Science and Technology Commission of Shanghai Municipality 24DP1500300.

Conflicts of Interest

None of the authors have a conflict of interest to disclose.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. J.-K. Han, D.-M. Geum, M. W. Lee, et al., "Bioinspired Photoresponsive Single Transistor Neuron for a Neuromorphic Visual System," *Nano Letters* 20, no. 12 (2020): 8781–8788, https://doi.org/10.1021/acs.nanolett.0c03652.

2. J. Chen, Z. Zhou, B. J. Kim, et al., "Optoelectronic graded neurons for bioinspired in-sensor motion perception," *Nature Nanotechnology* 18, no. 8 (2023): 882–888, https://doi.org/10.1038/s41565-023-01379-2.

3. X. Liu, D. Li, Y. Wang, D. Yang, and X. Pi, "Flexible optoelectronic synaptic transistors for neuromorphic visual systems," *APL Machine Learning* 1, no. 3 (2023): 031501, https://doi.org/10.1063/5.0163926.

4. X. Pang, Y. Wang, Y. Zhu, et al., "Non-volatile rippled-assisted optoelectronic array for all-day motion detection and recognition," *Nature Communications* 15, no. 1 (2024): 1613, https://doi.org/10.1038/s41467-024-46050-z.

5. T. Xie, Y.-B. Leng, T. Sun, et al., "Drosophila Visual System Inspired Ambipolar OFET for Motion Detection," *Advanced Functional Materials* 35, no. 7 (2025): 2415457, https://doi.org/10.1002/adfm.202415457.

6. J. Wei, G. Ma, R. Liang, et al., "Ferroelectric Optoelectronic Sensor for Intelligent Flame Detection and In-Sensor Motion Perception," *Nano-Micro Letters* 18, no. 1 (2026): 123, https://doi.org/10.1007/s40820-025-01968-x.

7. Z. Zhang, S. Wang, C. Liu, R. Xie, W. Hu, and P. Zhou, "All-in-one two-dimensional retinomorphic hardware device for motion detection and recognition," *Nature Nanotechnology* 17, no. 1 (2022): 27–32, https://doi.org/10.1038/s41565-021-01003-1.

8. J. Li, Z. Feng, Y. Qian, et al., "Organic Semiconductor Heterojunction-Based Photonic Synaptic Transistor with Temperature Stability and Adaptive Learning Ability for Neuromorphic Computing Systems," *ACS Applied Electronic Materials* 7, no. 13 (2025): 6149–6157, https://doi.org/10.1021/acsaelm.5c00882.

9. S. Jang, K. Soh, C. Lee, et al., "A unipolar-driven synaptic transistor for environment-adaptable vision system," *Nature Communications* 16, no. 1 (2025): 7636, https://doi.org/10.1038/s41467-025-63073-2.

10. X. Zhu, C. Gao, Y. Ren, et al., "High-Contrast Bidirectional Optoelectronic Synapses based on 2D Molecular Crystal Heterojunctions for Motion Detection," *Advanced Materials* 35, no. 24 (2023): 2301468, https://doi.org/10.1002/adma.202301468.

11. P. Xie, Y. Xu, J. Wang, et al., "Birdlike broadband neuromorphic visual sensor arrays for fusion imaging," *Nature Communications* 15, no. 1 (2024): 8298, https://doi.org/10.1038/s41467-024-52563-4.

12. H. Qu, H. Du, Y. Zhu, et al., "Voltage-Controlled Reconfigurable Boolean Logic Enabled by a β-Ga_2O_3 Artificial Optoelectronic Synapse," *ACS Applied Electronic Materials* 7, no. 21 (2025): 9792–9800, https://doi.org/10.1021/acsaelm.5c01613.

13. Z. Qu, Y. Xie, T. Zhao, et al., "Extremely Low Thermal Resistance of β-Ga_2O_3 MOSFETs by Co-integrated Design of Substrate Engineering and Device Packaging," *ACS Applied Materials & Interfaces* 16, no. 42 (2024): 57816–57823, https://doi.org/10.1021/acsami.4c08074.

14. E. S. Jo and Y. S. Rim, "Assessment of trapping layer control in IGZO/Al_2O_3/Ga_2O_3 synaptic transistor for neuromorphic computing," *Materials Today Physics* 37 (2023): 101194, https://doi.org/10.1016/j.mtphys.2023.101194.

15. J. Chen, H. Du, H. Qu, et al., "AlGaN/GaN MOS-HEMT enabled optoelectronic artificial synaptic devices for neuromorphic computing," *APL Machine Learning* 2, no. 2 (2024): 026113, https://doi.org/10.1063/5.0194083.

16. F. Gao, T. Wang, J. Bai, et al., "Amorphous Ga_2O_3 Based Erasable Two-Terminal Artificial Photonic Synapses," *Advanced Optical Materials* 13, no. 12 (2025): 2403227, https://doi.org/10.1002/adom.202403227.

17. H. Yoon, S. Park, Y. K. Kim, et al., "A Van der Waals Optoelectronic Synapse with Tunable Positive and Negative Post-Synaptic Current for Highly Accurate Spiking Neural Networks," *Advanced Functional Materials* 36, no. 16 (2025): 19498, https://doi.org/10.1002/adfm.202519498.

18. M. Zhou, Y. Zhao, X. Gu, et al., "Realize low-power artificial photonic synapse based on (Al, Ga)N nanowire/graphene heterojunction for neuromorphic computing," *APL Photonics* 8, no. 7 (2023): 076107, https://doi.org/10.1063/5.0152156.

19. S. Seo, S.-H. Jo, S. Kim, et al., "Artificial optic-neural synapse for colored and color-mixed pattern recognition," *Nature Communications* 9, no. 1 (2018): 5106, https://doi.org/10.1038/s41467-018-07572-5.

20. X. Chen, W. Mi, M. Li, et al., "Optoelectronic artificial synapses based on β-Ga_2O_3 films by RF magnetron sputtering," *Vacuum* 192 (2021): 110422, https://doi.org/10.1016/j.vacuum.2021.110422.

21. C. Kai, Y. Wang, X. Liu, et al., "AlGaN/GaN-Based Optoelectronic Synaptic Devices for Neuromorphic Computing," *Advanced Optical Materials* 11, no. 7 (2023): 2202105, https://doi.org/10.1002/adom.202202105.

22. Z. Long, Y. Ding, S. Poddar, L. Gu, Q. Zhang, and Z. Fan, "Bio-inspired visual systems based on curved image sensors and synaptic devices," *Materials Today Electronics* 6 (2023): 100071, https://doi.org/10.1016/j.mtelec.2023.100071.

23. Y. Zhang, Y. Huang, C. Huang, et al., "Plasma-Engineered AlGaN/GaN Optoelectronic Synapses for Low-Power Neuromorphic

Computing," *Advanced Optical Materials* 13, no. 13 (2025): 2403370, https://doi.org/10.1002/adom.202403370.

24. L. Yin, C. Han, Q. Zhang, et al., "Synaptic silicon-nanocrystal phototransistors for neuromorphic computing," *Nano Energy* 63 (2019): 103859, https://doi.org/10.1016/j.nanoen.2019.103859.

25. W. Fan, H. Yan, X. Wang, et al., "Polarization-Sensitive Photosynapse Based on $PdSe_2/WS_2$ Heterostructure for Visible-Infrared Broadband Artificial Vision System," *Advanced Functional Materials* 35, no. 43 (2025): 2416703, https://doi.org/10.1002/adfm.202416703.

26. Y. Zhang, L. Wang, M. Yang, et al., "Multiposition Controllable Gate WS_2/MoS_2 Heterojunction Phototransistor and Its Applications in Optoelectronic Logic Operation and Emulation of Neurotransmission," *Advanced Optical Materials* 10, no. 12 (2022): 2200197, https://doi.org/10.1002/adom.202200197.

27. S. Shrivastava, L. B. Keong, S. Pratik, A. S. Lin, and T.-Y. Tseng, "Fully Photon Controlled Synaptic Memristor for Neuro-Inspired Computing," *Advanced Electronic Materials* 9, no. 3 (2023): 2201093, https://doi.org/10.1002/aelm.202201093.

28. C. Liu, Y. Wang, W. Xu, et al., "Unique Bias Stress Instability of Heterogeneous Ga_2O_3-on-SiC MOSFET," *IEEE Electron Device Letters* 44, no. 8 (2023): 1256–1259, https://doi.org/10.1109/LED.2023.3288820.

29. C. Liu, B. Li, Y. Wang, et al., "Unique Monotonic Positive Shifts in Threshold Voltages of Ga_2O_3-on-SiC MOSFETs Under Both Unipolar Positive and Negative Bias Stresses," *IEEE Transactions on Electron Devices* 72, no. 3 (2025): 1047–1052, https://doi.org/10.1109/TED.2025.3534742.

30. W. Wang, E. Covi, A. Milozzi, et al., "Neuromorphic Motion Detection and Orientation Selectivity by Volatile Resistive Switching Memories," *Advanced Intelligent System* 3, no. 4 (2021): 2000224, https://doi.org/10.1002/aisy.202000224.

31. H. Wang, T. Liu, Y. Pan, et al., "Single-frame motion vector imaging with phonon-engineered perovskite optoelectronic synapses for neuromorphic visual perception," *Science Bulletin* 71, no. 7 (2026): 1612–1616, https://doi.org/10.1016/j.scib.2026.01.047.

32. Q. Yang, J. Hu, H. Li, et al., "All-Optical Modulation Photodetectors Based on the CdS/Graphene/Ge Sandwich Structures for Integrated Sensing-Computing," *Advancement of Science* 12, no. 11 (2025): 2413662, https://doi.org/10.1002/advs.202413662.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting file: adom71653-sup-0001-SuppMat.docx.