REVIEW ARTICLE

WILEY

Next-generation machine vision systems incorporating two-dimensional materials: Progress and perspectives

Peisong Wu[1,2] | Ting He[1,3] | He Zhu[3] | Yang Wang[1,4] | Qing Li[3] | Zhen Wang[1,2] | Xiao Fu[3] | Fang Wang[1] | Peng Wang[1] | Chongxin Shan[4,5] | Zhiyong Fan[6] | Lei Liao[7] | Peng Zhou[4] | Weida Hu[1,2,3]

[1]Key Laboratory of Infrared Physics, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China

[4]State Key Laboratory of ASIC and System, School of Microelectronics, Fudan University, Shanghai, China

[5]Henan Key Laboratory of Diamond Optoelectronic Materials and Devices, School of Physics and Engineering, Zhengzhou University, Zhengzhou, China

[6]Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China

[7]School of Physics and Electronics, Hunan University, Changsha, China

Correspondence
Peng Zhou, State Key Laboratory of ASIC and System, School of Microelectronics, Fudan University, Shanghai 200433, China.
Email: pengzhou@fudan.edu.cn

Weida Hu, Key Laboratory of Infrared Physics, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China.
Email: wdhu@mail.sitp.ac.cn

Abstract

Machine vision systems (MVSs) are an important component of intelligent systems, such as autonomous vehicles and robots. However, with the continuous increase in data and new application scenarios, new requirements are put forward for the next generation of MVS. There is an urgent need to find new material systems to complement the existing semiconductor technology based on thin-film materials, and new architectures must be explored to improve efficiency. Because of their unique physical properties, two-dimensional (2D) materials have received extensive attention for use in MVSs, especially in biomimetic ones: the human visual system, which can process complex visual information with low power consumption, provides a model for next-generation MVSs. This review paper summarizes the progress and challenges of applying 2D material photodetectors in sense-memory-computational integration and biomimetic image sensors for machine vision.

KEYWORDS

in-sensor computing, machine vision systems, photodetectors, two-dimensional materials

1 | INTRODUCTION

Machine vision systems (MVSs) simulate human visual functions to observe and recognize the objective world.[1,2] They can be used for image acquisition, image processing, and feature recognition.[3] With the rapid development of information integration, machine vision has become indispensable in automated production, autonomous driving, robotics, security, and other domains.[2] New application scenarios and increasing quantities of data have created a demand for MVSs with faster parallel processing, higher energy efficiency, smaller volume, and lower price.[4]

Image acquisition by an MVS depends on the image sensor. At present, most commercial visible-image sensors are either charge-coupled devices (CCDs) or CMOS image sensors (CISs).[3] Infrared photosensor materials include cooled mercury-cadmium-telluride (MCT), InSb, quantum wells, Type-II superlattice or uncooled VO_x, PbS, InGaAs, amorphous silicon, PbS, and PbSe.[5] In recent years, two-dimensional (2D) materials (i.e., materials with atomically thin layers) have become a powerful competitor.[6–10] Their extreme thinness, quantum electronic states, large carrier mobility, and localized optical transition give them many advantages over traditional materials.[11] 2D materials are a large family, including metals, semi-metals, semiconductors, and insulators.[12] Because they are so thin, their band structure can be adjusted through a localized electric field or external stress; thus, their detection waveband can range from ultraviolet to terahertz.[13] Because different layers of 2D materials are bonded by van der Waals forces, with no dangling bonds, they can be easily integrated with each other or memristor circuits, regardless of lattice mismatch or deposition temperature.[11,14]

Combining 2D materials with silicon chips to form a heterogeneous structure platform based on vdW forces can greatly promote vertical integration and functional diversification.[15,16] However, an MVS based on von Neumann architecture encounters difficulties when applied to new scenarios with higher performance requirements. The image acquisition, storage, and processing modules in traditional von Neumann structures are physically separated.[17] Visual information is obtained through image sensors in the MVS and stored in a storage module. To conduct further image processing requires that a large amount of image data be transferred repeatedly between the memory and the processor, resulting in high energy consumption.[18] In addition, the bandwidth of the storage module is much smaller than that of the processing module: the so-called "storage-wall problem", which limits the speed of the entire system.[19] Therefore, when implementing parallel tasks such as image processing, the von Neumann structure has low efficiency, high energy consumption, and is slow.[2,20]

The human visual system (HVS) can recognize various objects and perceive visual information in a complex environment with high energy efficiency.[21] Image information is preprocessed near the retina before being sent to the brain.[22] Therefore, there is no need for an analog-to-digital (AD) conversion process, which would require high power consumption. HVS-inspired machine vision preprocessing functions utilizing a 2D-material photodetector have therefore been proposed. To improve data processing capacity, speed, and power consumption further, neural networks can be combined with such novel photodetectors. A 2D-material photodetector array can be integrated with a memristor module, the resistance value of which can then be used as the weight of the neural network. More importantly, because of the 2D material's unique layered structure, the carrier density can be precisely tuned by shifting the gate voltage, and even the carrier type can be changed, altering the response range or responsivity.[23,24] The adjustable responsivity can mimic the variable weights of a neural network, allowing preprocessing to occur in the detection module and thereby increasing speed and reducing energy consumption. Detection, calculation, and storage functions can be realized in one device simultaneously.[9]

This review summarizes a variety of attempts to apply 2D-material photodetectors to MVSs, including vertical integration with silicon-based readout circuits and integrated sensing, memory, and computing architecture. As a new material system, 2D materials face many challenges in the transition from the laboratory to industry. This paper considers not only the advantages and diversified functions of 2D materials but also the limitations of existing MVSs. Some possible future applications are considered in detail. This review provides a basis for the application of 2D material photodetectors in terms of materials, performance, and systems. The main challenges to practical applications are highlighted, including the preparation of large-area materials and the calibration of key performance indicators. We suggest a possible solution to the latter problem, at least for the case of normalized detectivity.

2 | VERTICALLY INTEGRATED MVSs BASED ON 2D-MATERIAL PHOTODETECTORS

2.1 | Vertical integration with silicon-based von Neumann architecture

Silicon semiconductor technology, the foundation of the modern electronics industry, has been developing vigorously since the 1960s. The appearance of 2D materials

makes possible future development of the microelectronics industry at the atomic scale without considering lattice mismatch. Benefitting from van der Waals forces between layers, 2D materials can be stacked like LEGO blocks at the atomic level without restrictions on lattice matching. Therefore, 2D-material-based detectors can be vertically integrated with Si-based chips at high densities (Figure 1C). Already, a 5 × 5 MoS$_2$-based detector array has been heterogeneously integrated with a silicon chip[26] (Figure 1D). In another study, a 388 × 288 detector array using graphene and quantum dots was heterogeneously integrated with a CMOS readout circuit with a detection waveband of 2000 nm[25] (Figure 1A). This was the first time that high-density vertical interconnection between an infrared detector and a silicon-based chip was realized. Graphene was used as the channel of the photodetector, while PbS quantum dots were spin-coated on graphene as the absorption layer. The working mechanism of this device was photogating. The bottom silicon chip, acting as the readout circuit, integrated and amplified the signal from the detector (Figure 1B). This type of heterostructure fabricated by dry transfer[27,28] is particularly important in the development of infrared detection, because it is not restricted by lattice matching between materials.

As described by Liu et al.,[11] vdW integration can be further developed. The detector, memory, and logic circuit can be vertically integrated to realize a system-level vdW detection chip (Figure 1E). This technology significantly reduces the MVS volume while increasing the pixel density. However, it is still based on the von Neumann structure and is subject to the storage wall and other problems that have not yet been solved. Future MVS systems need new architectures integrating detection, storage, and processing capabilities at the device level.

2.2 | Limitations of von Neumann architecture

In existing MVSs, light passes through a set of lenses before converging onto photodetectors, which convert optical signals into analog photocurrent signals. The readout circuit integrates the analog photocurrent signal into an analog voltage. After filtering and noise reduction, the analog voltage is converted into a digital signal by AD conversion.[29] The central processing unit (CPU) or graphics processing unit (GPU) conducts noise reduction, nonuniformity correction, graphics distortion correction, and other processing before sending the digital signal into a trained neural network for recognition.[29,30] The working process of the MVS is illustrated in Figure 2A. The architecture is inefficient
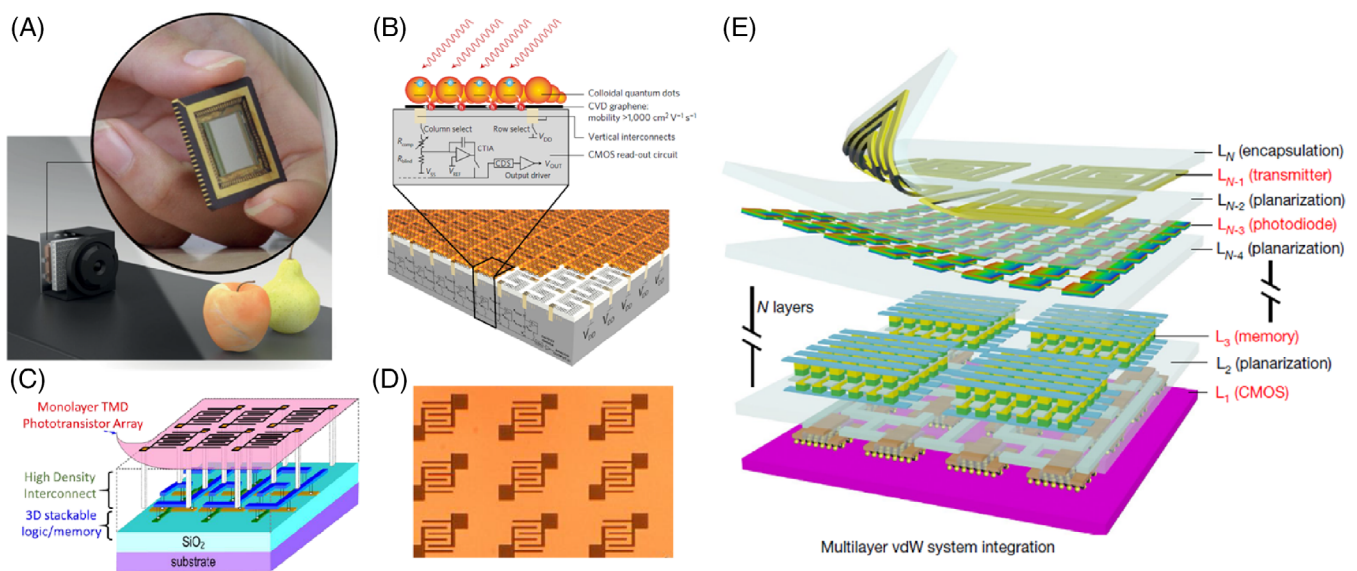

FIGURE 1   (A) Imaging system based on CMOS-integrated graphene-quantum dot photodetectors.[16] Inset: a photograph of the packaged monolithically integrated graphene-based image sensor. (B) Side view explaining the graphene photoconductor and the underlying readout circuit.[25] (C) Schematic illustration of a 5 × 5 MoS$_2$-based detector array and silicon chip.[26] (D) Image of the top monolayer TMD photodetector array.[26] (E) Detector, memory, and logic circuits can be vertically integrated to realize the level detection chip of vdW system.[11] (A) Reproduced with permission.[16] Copyright 2019, Springer Nature. (B) Reproduced with permission.[25] Copyright 2017, Springer Nature. (C,D) Reproduced with permission.[26] Copyright 2016, IEEE. E, Reproduced with permission.[11] Copyright 2019, Springer Nature

in terms of power consumption and data-transfer speed, in that the memory and processing units are physically separated.[31]

A near-sensor architecture adopting near-memory computing has been proposed that shortens the transmission distance between storage and computation and improves energy efficiency in traditional-film-material systems.[32] Near-memory computing frameworks have been based on cutting-edge packaging technologies, such as system-in-package (SiP), 2.5D SiP, and 3D SiP technologies.[33] There has been no attempt to introduce system-on-a-chip (SoC) technology (which integrates CPU, dynamic random-access memory (DRAM), read-only memory (ROM), input/output interface (IO), analog-to-digital converter (ADC), and other modules on a chip) into image-sensor design, because the CIS and CPU use different fabrication processes. In particular, most infrared-sensitive materials, such as InGaAs, InSb, MCT, and VO_x, are not easily compatible with the CMOS process. One popular method of 2.5D/3D packaging, through silicon via (TSV), encapsulates chips with different functions on one substrate, thereby significantly shortening the length of the connection between different modules and improving the connection density[34] (Figure 2C). Sony has widely adopted

TSV packaging technology in visible CIS designs[35,36] (Figure 2B). In addition, by adopting the TSV package, the newly released infrared InGaAs product significantly reduces pixel size and improves pixel density and transmission speed, resulting in a more colorful and higher frame-frequency imaging effect. However, this technology is still not mature enough; it has a high price and low yield compared with other packaging technologies. TSV packaging for other infrared-sensitive materials is at the laboratory prototype stage, and many technical problems still need to be solved. Taking MCT as an example, defects introduced by the TSV process result in a high dark current and other problems.

In essence, near-memory computing is still dominated by the von Neumann architecture. In the 10 nm process, the power consumption of data bus accounted for more than 69% (Figure 2E). Its transmission speed is limited by the connection density and will soon hit a bottleneck. The speed of the existing high-performance photodetector is less than 1 ns; the response speed of the ROM is on the order of milliseconds. Thus, the storage speed is much lower than the detection speed. The speed mismatch between the modules also significantly affects the recognition speed (Figure 2D).
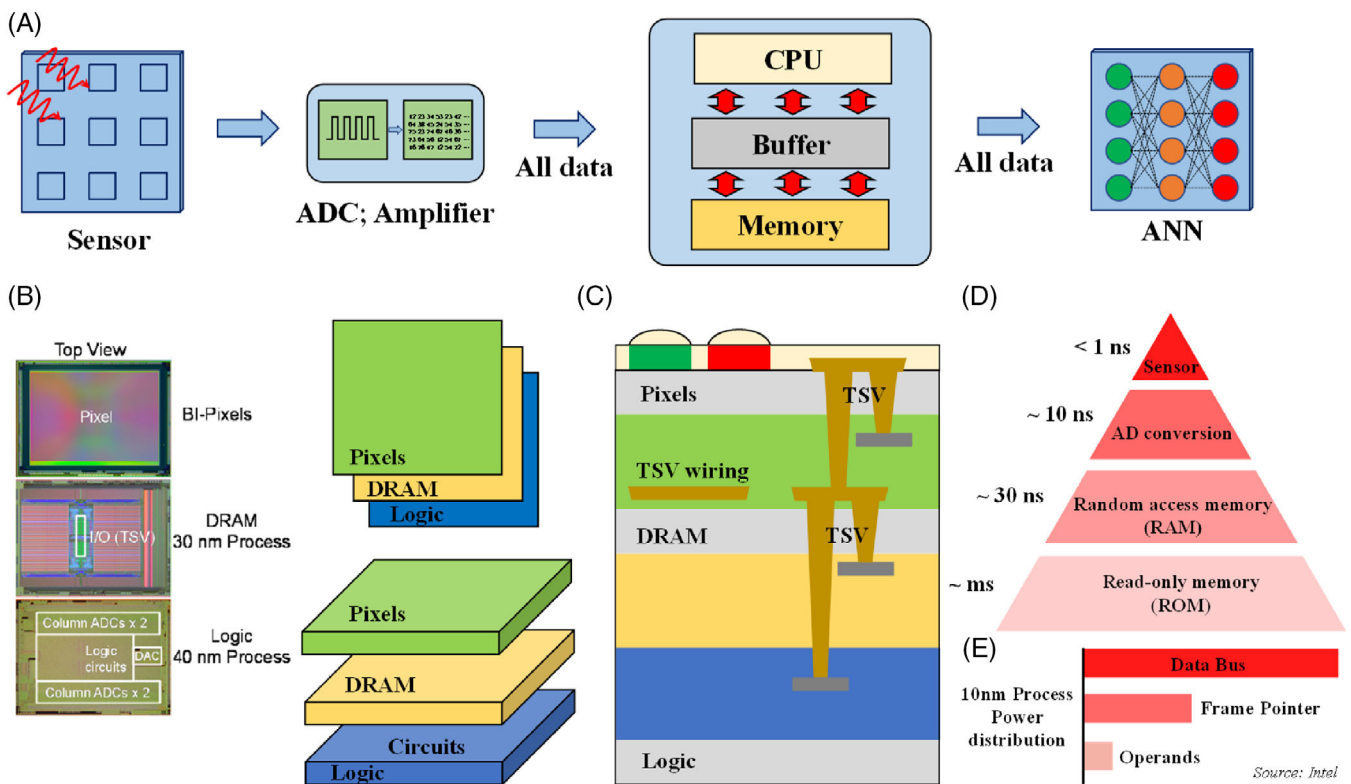

FIGURE 2 (A) Working process of the MVS system. (B) Schematic diagrams of three-layer stacked photodetector. (C) Schematic of cross-section of three-layer stacked sensor profile. (D) Schematic diagram of speed comparison between modules. (E) Schematic diagram of energy per compute operation

3 | NEW ARCHITECTURE FOR MACHINE VISION SYSTEMS

3.1 | The human visual system

Nature has been a source of ideas for engineers in such diverse fields as aerodynamics, robotics, surface, structural engineering, and materials science.[37–39] The natural HVS was developed through a long evolutionary process and may provide inspiration for solving the problems faced by MVS. It can recognize various objects and perceive visual information in a complex environment with high energy efficiency.[21]

The HVS has two main advantages over existing MVSs. First, human eyes do not perform AD conversion after receiving visual information, but directly preprocess the analog signal in the retina, greatly improving the processing speed while maintaining high-enough computational accuracy for the eye to be a major source of information about the outside world for most humans.[40] In addition, image information is preprocessed near the retina before sending it to the brain, which increases the possibility of recognition and energy efficiency.[22]

The human retina is composed of a photoreceptor layer, bipolar cell layer, and ganglion cell layer. The light signal is first received by photoreceptors, which are divided into cones and rods (Figure 3A). Color determination and light collection under strong illumination are performed by cone cells, light collection under weak illumination by rod cells. The signals collected by the photoreceptors are then transmitted to the bipolar cells to generate bidirectional electrical signals that are modified by amacrine cells before being transmitted further.[41] Amacrine cells not only modify the output from the bipolar cells but also that from other associated amacrine cells, so different combinations of them perform different functions. Then, the modified information is sent to the ganglion cells (and eventually to the brain).[42]

For initial integration or preconditioning in the retina, the electrical signal goes to the excitation-inhibition neural network, which is composed of intermediate nerve cells; this network realizes the center-surround antagonism function and improves the ability of ganglion cells to recognize contours.[40] The central peripheral antagonistic receptive field of ganglion cells corresponds to the difference of Gaussians in the convolutional neural network (CNN) algorithm (Figure 3B). It can convolve the original input image to extract the outline of the image (Figure 3C). Next, the retina categorizes and packages preprocessed information into the optic chiasm and further into the visual cortex of the brain.

The human brain is composed of many neurons connected through synapses; it implements learning and complex cognitive tasks with high energy efficiency, playing the role of a CPU. The BNN in the human brain consists of neurons rich in synapses that undertake the responsibilities of computation and storage simultaneously.[43] Thus, mimicking the HVS to integrate sensing, storage, and processing could solve the problem of speed mismatch and low energy efficiency in the von Neumann architecture.[44]
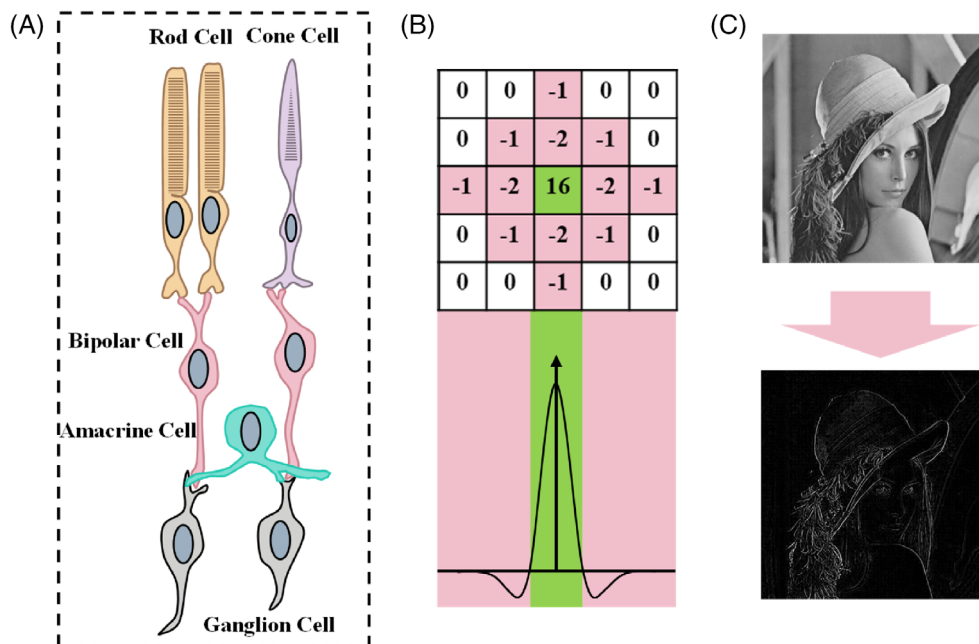

FIGURE 3 (A) Profile of a biological retina. (B) Schematic of central peripheral antagonistic receptive field in ganglion cells. (C) Original Lenna image. The processed image by edge enhancement

3.2 | In-sensor preprocessing for machine vision systems

The human retina can perform many preprocessing steps, including integrating visual information and extracting contour information, that greatly reduce the processing load of the brain and improve energy efficiency.[45–47] Photoelectric hybrid neural networks based on photonic memristors can mimic the sensing and preprocessing functions of HVS (Figure 4A). Zhou et al. demonstrated a two-terminal photonic memristor[48] (Figure 4B). This device had optically tunable characteristics, including long-term memory, short-term memory, and a tunable storage time affected by the written optical-pulse dose.

The nonlinear change in the resistance could function as an HDR (high-dynamic range) preprocess in image processing. The main features of the image were highlighted after preprocessing, and the image contrast was enhanced. The recognition efficiency of the algorithm was improved compared with that of the neural network without image preprocessing: after 1000 training epochs, the recognition rate with image preprocessing reached 0.986, while that without image preprocessing was only 0.980 after 2000 training epochs (Figure 4C). Photonic memristor arrays can realize image sensing and memory functions, as well as neural morphological vision preprocessing. The proof-of-concept device improves the processing efficiency and image recognition
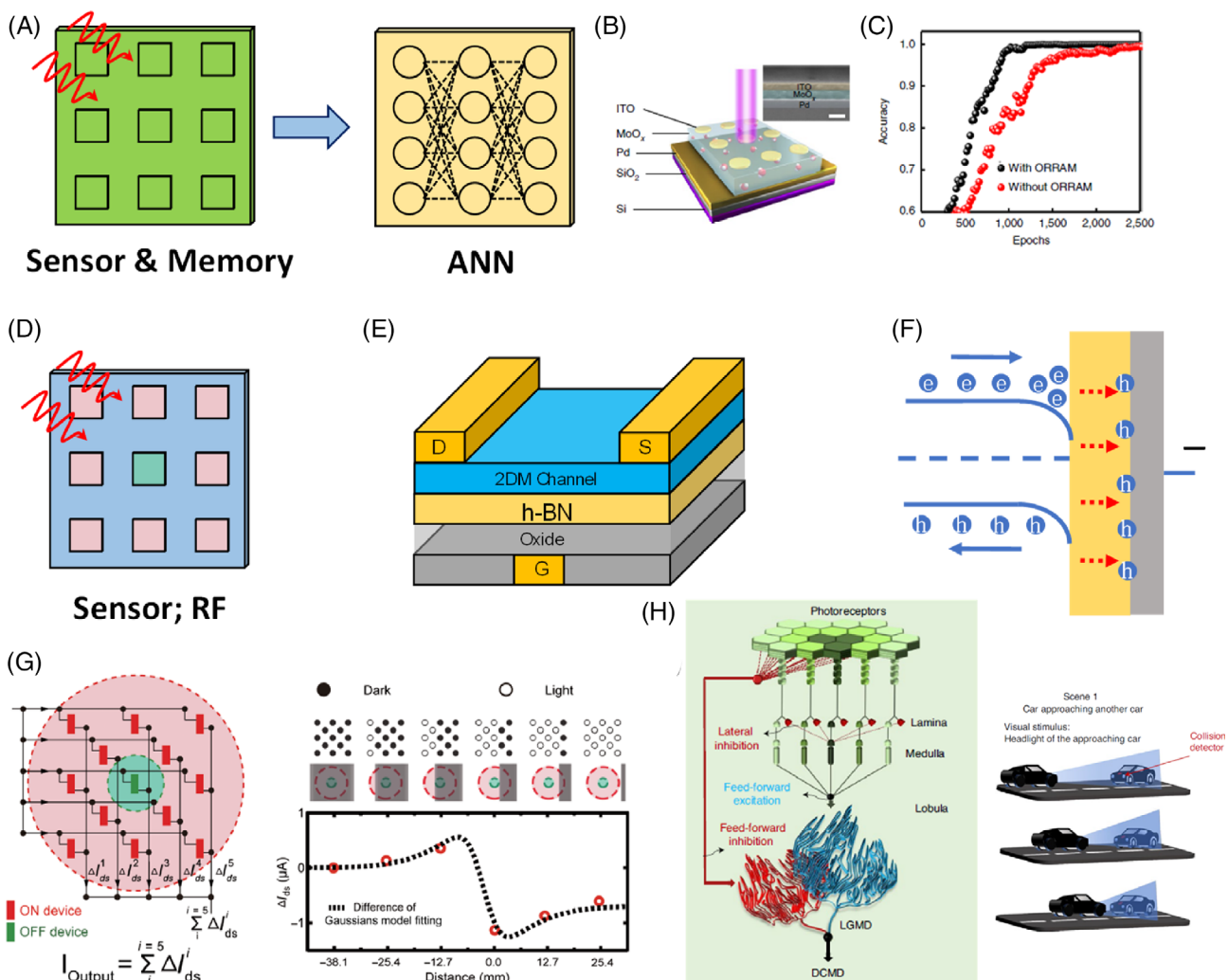

FIGURE 4 (A) Photoelectric hybrid neural networks based on photonic memristor. (B) Schematic structure of the MoO_x photonic memristor.[48] (C) Preprocessing shortens the training cycle.[48] (D) Sensor mimics the biological receptive field to realize the preprocessing function. (E,F) Floating-gate structure can adjust channel responsiveness and realize storage function. (G) Schematic diagram of RF preprocessing.[40] Light panel: Sensor with preprocessing function and its circuit configuration. Right panel: The photocurrent changes dramatically at the edges of the image. (H) A schematic diagram of the locust visual system, illustrating the neurobiological anti-collision architecture.[39] (B,C) Reproduced with permission.[48] Copyright 2019, Springer Nature. (G) Reproduced with permission.[40] Copyright 2020, American Association for the Advancement of Science. (H) Reproduced with permission.[39] Copyright 2020, Springer Nature

rate, which contribute to the development of edge computing.

MSV researchers have also attempted to imitate the type of pattern recognition that occurs in the retina of the human eye.[40] Bipolar cells connected to photoreceptors are classified into ON and OFF types that respond to light stimulation oppositely. This architecture, which enables the human eye to preprocess the image and transmit it to the cerebral cortex for further processing (Figure 4D), is analogous to the use of floating gates to regulate the responsivity of an artificial detector. Unlike a traditional FET, a floating gate adds a charge-binding layer to the dielectric layer; the charge of the floating-gate layer is controlled by the impulse voltage on the control gate (Figure 4E,F). Thus, the carrier concentration in the channel is controllable: photodetectors based on floating-gate memory could integrate detection and storage by adjusting the channel conductance.[49]

Wang et al. demonstrated a bio-inspired WSe_2/h-BN/Al_2O_3 vertical heterostructure that mimics the vertical integration of the photoreceptor and bipolar cell layers in the retina.[40] This device could be regulated by the back-gate voltage to achieve a positive or negative photoresponse imitating that of the bipolar cell layer (Figure 4G). An OFF-device surrounded by 12 ON-devices could mimic the biological receptive field in the HVS that detects the edges of objects. In the experiment, the light-source array transmitting the image was turned on column by column, imitating a moving edge. As the edge moved from the left side to the right, the current increased because of the activation of more ON-devices. It reached a maximum value before the edge reached the central OFF-device, which was later activated to produce a reverse photocurrent. The entire photocurrent output of the system decreased, resulting in an opposite photocurrent peak. The sharp change in photocurrent realized the recognition of the image boundary. Another task was image recognition without AD conversion. The central peripheral antagonistic receptive field formed by the array performed a convolution operation on the original input image to extract the contour. The CNN formed by the prototype device was used to classify the target image.

Jayachandran et al. created an extremely low-power, dynamic, nonvolatile collision-avoidance monitoring system by vertically stacking a single-layer MoS_2 light detector with a floating-gate transistor-based memory device to simulate the LGMD neurons in locusts[39] (Figure 4H). These are unique and complex neurons that use a distributed computing architecture to assist in processing visual information. The honeycomb-shaped photoreceptors in the insect's compound eye convert visual stimuli into electrical impulses that pass through the lamina, medulla, and lobules, before eventually being transmitted to the dendritic fan-out area of LGMD neurons. The dendritic branches in the red area receive feedforward suppression; those in the blue area receive feedforward excitation and lateral suppression. The system developed by Jayachandran et al. used positive feedback from the light detector to the light signal and negative feedback from the floating gate to simulate this natural computational architecture. The photodetector responded to the impending object (stimulus signal) and caused the device current to increase, while the underlying programmable memory stack always caused the current (inhibition signal) to decrease. When an object approached, an excitatory signal was added to the inhibitory stimulus; this caused a non-monotonic change in the device current, simulating the escape response of LGMD neurons in the locust. Thus, the system could detect an impending collision in time and trigger an escape response with nanojoule energy consumption. This in-memory task-specific computation and perception method preprocessed the raw data, greatly simplifying the computation and reducing the energy consumption.

3.3 | In-sensor computing for machine vision systems

Because of their vdW bonding, 2D materials allow arbitrary stacking to form 3D high-density integration with an electronic memristor array. Connecting 2D photodetectors with a memristor array is an effective way to improve energy efficiency. A memristor is a type of non-volatile memory, the resistance value of which can be changed nonlinearly by an external stimulus[48,50,51] (Figure 5A). The adjustable resistance can be used as the weight of the neural network. It can simultaneously process and store information that would require a dozen CMOS transistors in matrix computing.

Wang et al. proposed a neuromorphic visual system composed of sensors and memory networks mimicking the hierarchical organization and biological functions of the retina.[4] Figure 5B shows a flowchart and processing of the neuromorphic visual system. The sensor was constructed using a WSe_2/h-BN/Al_2O_3 vdW heterostructure detector with the ability to adjust electrically between ON and OFF states. Some common kernel functions in a CNN algorithm were realized by adjusting different detectors into the ON/OFF state in a 3×3 bionic retinal sensor (Table 1 summarizes commonly used convolution kernels and their corresponding ON/OFF states). A variety of image-processing operations, such as useless-information filtering and key-information extraction, were realized using different convolution kernels. However, the kernel function represented by the sensor was not adjustable during its operation, unlike the dynamic kernel function in a normal CNN algorithm. Moreover,
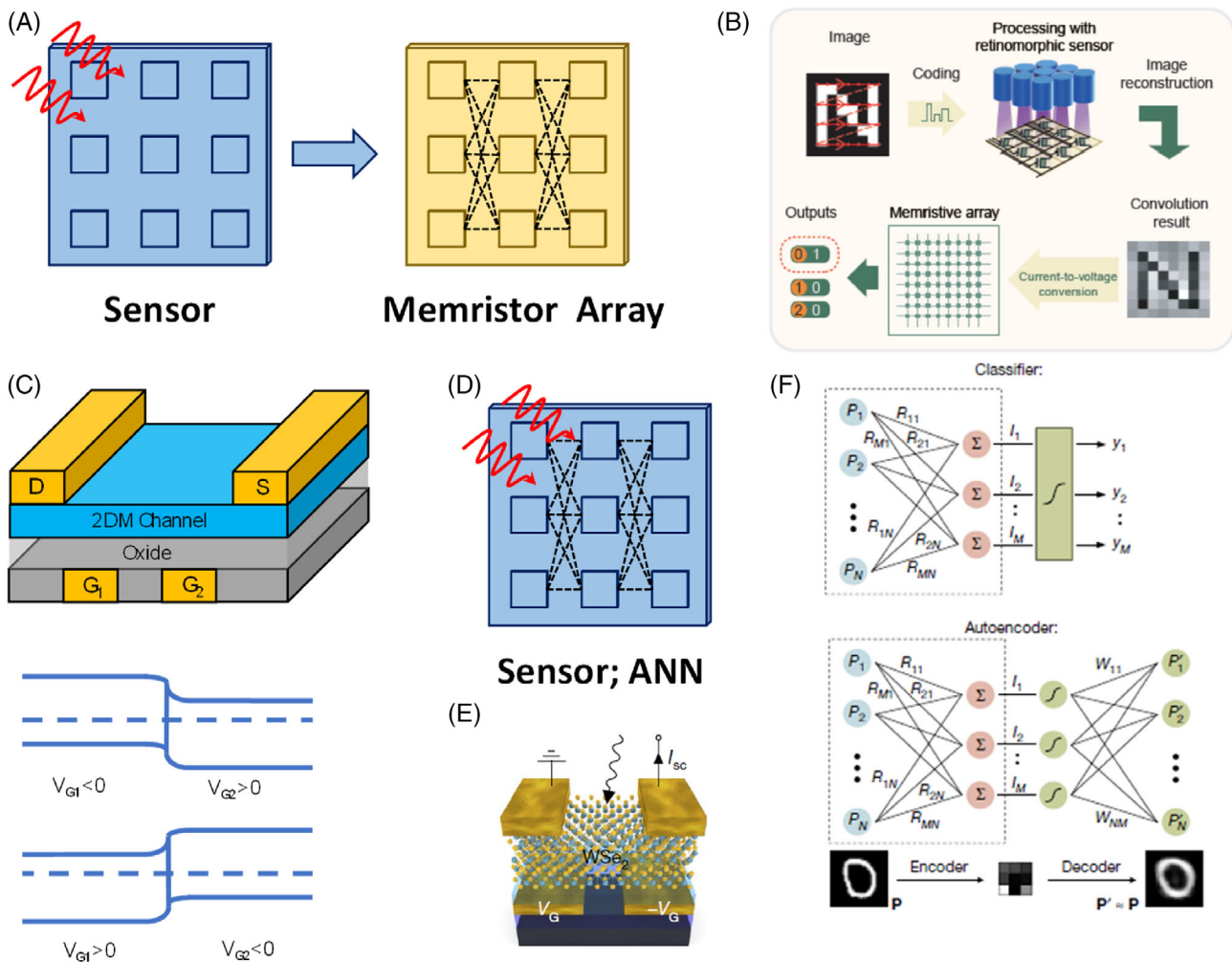
FIGURE 5 (A) Connecting 2D photodetectors with the memristor array is an effective way to improve energy efficiency. (B) Flowchart briefly illustrates the sensing and processing of images by a detector combined with a memristor array.[4] (C) 2DM photodetector with adjustable response rate is realized by adjusting 2DM band with double back gates. (D–F) Mennel et al. reported a system that combines a sensor chip with a neural network: (D) Sensor can not only collect signals but also act as a neural network to recognize simple features. (E) Schematic diagram of unit WSe₂ photodiode. The device works under short circuit condition and the photoresponse is set by the bottom gate electrode. (F) System can be trained in Classifier and Autoencoder two modes.[9] (B) Reproduced with permission.[4] Copyright © 2020, Oxford University Press. (E,F) Reproduced with permission.[9] Copyright 2020, Springer Nature Publishing AG

the extracted optoelectronic signal still required transformation into a digital signal before being output into the memristor crossbar network. Although the packaging was not specified in the paper, it seems more likely to have been PCB than SiP. In contrast to von Neumann structure-based processors, the memristor can process analog information directly without AD conversion. The neuromorphic system accomplished excellent cross-tracking by networking the retinomorphic sensor with a recurrent neural network in a fully analog signal environment.

Commercial silicon-based photodetectors show a fixed responsivity that is determined by the structure and doping concentration inside the device. The physical properties of 2D materials are easily regulated by external

stimuli because of their atomic layer thickness.[52] The applied voltage controls the carrier concentration and polarity of the 2D materials (Figure 5C). Because of the unique layered structure of 2D materials, the carrier density can be precisely tuned by shifting the gate voltage. Even the carrier type can be changed, resulting in a change in the response range or responsivity. A homogeneous WSe₂ P–N junction has been realized by regulating the local carrier density in different WSe₂ areas using a double gate.[24] Responsive-regulated regulated P–N diodes based on this homogeneous junction were obtained. The responsivity of the photodiodes could reach 210 mA W⁻¹, and the corresponding energy efficiency conversion was 0.2%. The adjustable responsivity could mimic the variable weights of a neural network. This means that the new

TABLE 1 Commonly used convolution kernels and the corresponding on/off states

Convolution kernels	ON/OFF states	Function
	Inverse	Image stylization
	Difference-of-Gaussian	Edge enhancement
	Laplacian	Contrast correction

Note: ■ ~ON ■ ~OFF ■ ~Zero

photoelectric detector can integrate the processing function into the sensing module, thereby improving visual processing speed and reducing power consumption.

If some physical parameters of the photodetector, such as the resistance of an electrical memristor, can be dynamically regulated by an external voltage, a neural network can be formed directly by the photodetector array. Mennel et al. designed photodiode arrays based on 2D materials, providing a new architecture to replace the traditional von Neumann architecture for image recognition. This enabled the implementation of more in situ calculations on the detector[9] (Figure 5D). As a photodiode array, the device itself constituted a neural network that could recognize and process images simultaneously, breaking through the limitations of traditional separation modules. This significantly improved the speed and efficiency of intelligent recognition. Double-gate electrodes were used to dope channel materials electrically, resulting in an adjustable responsivity for each pixel device (Figure 5E). Through the artificial neural network, the system executed supervised and unsupervised learning tasks (Figure 5F). The responsivity of each unit served as the weight of the neural network. The responsivity was constantly adjusted to maximize the system's identification accuracy. Because both image sensing and image processing were carried out in an analog signal environment, the speed of the system was limited only by the optical–electrical conversion process. Therefore, image detection and recognition occurred

within 50 ns. Theoretically, the system could process 20 million images per second, a rate several orders of magnitude faster than that of traditional methods. In addition, compared with current MVSs, which consume a large amount of power per operation, the biological neural network had very low energy consumption per operation: 10^{-15} to 10^{-13} J.

4 | CHALLENGES OF TWO-DIMENSIONAL MATERIAL ARRAY DETECTORS

Photodetectors using 2D materials may enable the development of novel photoelectric sensors for energy-efficient computing and in situ recognition and of curved image sensors with remarkable volume advantages, as will be discussed below. However, commercial viability has not yet been attained. Some technical issues still need to be addressed, including the stable growth and doping of large-scale crystals, standardization of evaluation methods, and development of large-scale heterogeneous integration techniques. In this section, we discuss the obstacles to material growth and device performance, as well as possible ways to overcome them.

4.1 | Large-area two-dimensional materials and photodetector arrays

Large-area materials are indispensable for large-scale integrated optoelectronic devices. Although small samples of 2D materials can be easily obtained by mechanical exfoliation, the growth of large-area high-quality 2D materials remains a major challenge. Because the thickness of the single layer of 2D material, unlike that of a conventional thin film, is of nearly atomic scale, growth depends on the substrate surface. Multilayer 2D materials, while lacking some of the advantages of single layers, have good absorption of light and are widely used in the field of photodetectors, but the van der Waals force between the layers is weak, and the growth process is difficult to control.[53] The growth method with a controllable number of layers is also a challenge currently faced.

Many factors may influence the growth of high-quality large-area 2D materials, such as elemental ratio, surface defects, nucleation processes, growth catalytic techniques, and phase control, require further investigation. Because of the immature state of growth technology, defects are produced in the material: vacancies, substitutions, anti-sites, and adsorbed atoms.[54] These have a significant influence on the electrical and optoelectronic properties of 2D materials: they can act as carrier donors,

scatterers, traps, and recombination centers under different conditions.[55] In the preparation process, it is necessary to avoid most defects, but sometimes beneficial defects can be used to achieve special functions. Therefore, a systematic understanding of 2D material growth is necessary for the future development of detector arrays. The construction of high-quality heterogeneous structures incorporating large-area materials is another challenge to be faced in the future.

Because of the stringent requirements for substrate quality in the growth process of 2D materials,[53] it is necessary to eliminate potential excess nucleation points before growth. In the early stage of growth, graphene has been grown on the surface of Cu foil, and the active parts of the foil surface have been passivated through high-temperature, long-time annealing[56] or electrochemical polishing[57] to achieve millimeter-sized growth. Oxidization of growing substrates to eliminate potential nucleation points has been widely used and intensively studied, pushing the size of graphene domains up to the centimeter scale.[58–61] By using Ni catalysis and CH_4 gas to induce epitaxial growth, it is even possible to prepare meter-scale single-crystal graphene.[62] However, graphene is a semiconductor with a zero bandgap. Graphene's poor absorption limits its applicability in photodetection. Therefore, graphene is generally used in heterostructures. Goossens et al. used large-area graphene and PbS quantum dots to form heterostructures and realized a 388×288 near-infrared camera[25] (Figure 6A). Quantum dots acted as the photosensitive layer, while graphene, because of its high mobility, became a carrier fast-transport channel.[25] When photoexcited, free electron–hole pairs were generated in the quantum dots; the holes migrated to the graphene, whereas the photogenerated electrons stayed in the quantum-dot layer, where they could potentially regulate the graphene channels.[63] As a conductive channel with high mobility, graphene is easily affected by an external electric field; as a photosensitive absorption layer, PbS quantum dots compensate for the weakness of weak light absorption and realize near-infrared detection by improving the gain of the photogating effect (Figure 6B).

Because of their excellent electronic and optical properties, transition-metal dichalcogenides (TMDs) are promising materials for next-generation electronic and optoelectronic devices. The large-area growth of TMDs can be achieved through chemical vapor deposition (CVD), depositing metal precursors before sulfurization or selenization. This can produce centimeter scale, atomically thin, and uniform MoS_2,[67,68] $NbSe_2$,[69] and $PdSe_2$[70] crystals. In addition, pulsed laser deposition (PLD),[71] atomic layer deposition (ALD),[72] and metal–organic chemical vapor deposition (MOCVD)[73] have been used to grow large-area materials. For example, in the case of
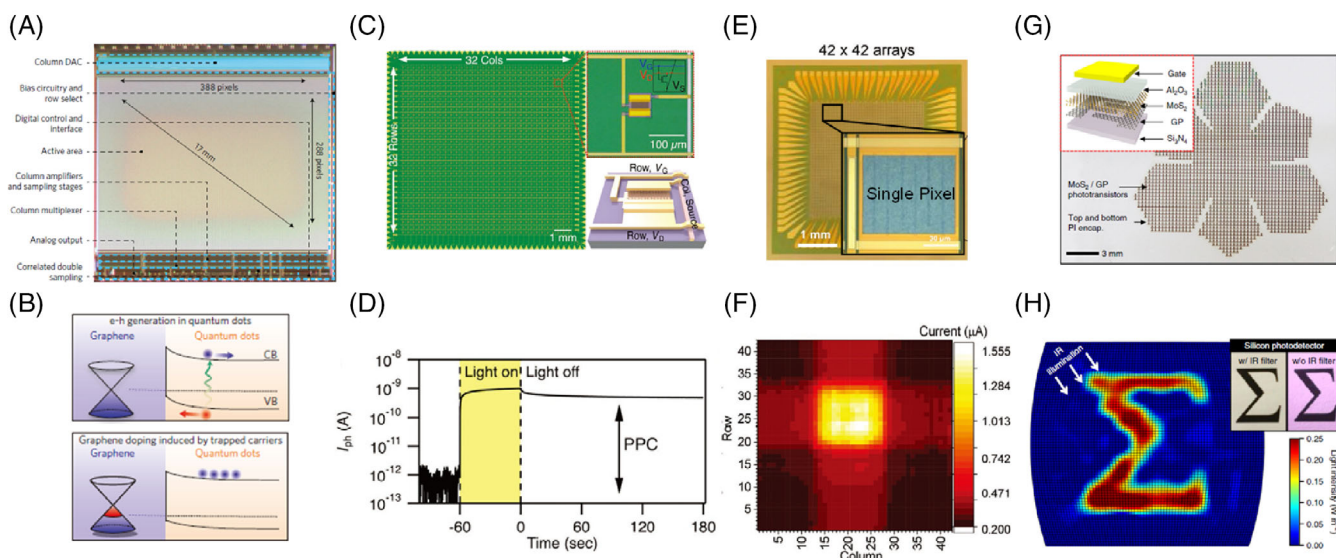
FIGURE 6 Two-dimensional material detector array. (A) Back-end-of-line CMOS integration of CVD graphene with 388×288 pixel image sensor readout circuit.[25] (B) Energy level diagram of the graphene/quantum dot interface. (C) Optical microscopy images of the 32×32 MoS_2 photo-FET crossbar array. (D) The persistent photoconductivity effect of 32×32 MoS_2 photo-FET. (E) 42×42 $SeTe_2$ photodetector array. (F) Imaging based on $SeTe_2$ focal plane detector array. (G) Schematic illustration of the device design. Inset shows a schematic illustration of the device structure. (H) Infrared imaging of the curved image sensor array. (A) Reproduced with permission.[25] Copyright 2017, Springer Nature Publishing AG. (B) Reproduced with permission.[63] Copyright 2012, Springer Nature Publishing AG. (C,D) Reproduced with permission.[64] Copyright 2020, John Wiley & Sons, Inc. (E,F) Reproduced with permission.[65] Copyright 2020, John Wiley & Sons, Inc. (G,H) Reproduced with permission.[66] Copyright 2017, Springer Nature Publishing AG

MoS$_2$, anti-parallel domains (0° and 60° orientations) appear during synthesis because of the special triple symmetry of the lattice, resulting in double grain boundaries during the splicing process.[74] Therefore, the main practical obstacle is the accurate control of the unidirectional arrangement of MoS$_2$ domains on the wafer.

The successful growth of large-area photosensitive materials has greatly promoted the development of detector arrays based on 2D materials. Yang prepared a 5 × 5 MoS$_2$ photodetector array with a responsivity of 25 A W^{-1} at 637 nm.[26] Jang prepared a 32 × 32 MoS$_2$ photodetector array[64] (Figure 6C) and realized the internal neural network calculation by using the continuous photoconductance of MoS$_2$ (Figure 6D); this system successfully recognized handwritten digits. Tan et al. fabricated a 42 × 42 Se$_x$Te$_{1-x}$ array as a 2D-material camera for infrared imaging[65] (Figure 6E,F).

Large-area 2D materials are also prone to forming heterojunctions because of vdW forces. Zeng et al. used controlled growth of PdSe$_2$ combined with Si to realize a 4 × 4 photodetector array, achieving highly sensitive detection performance at 780 nm.[70] An 8 × 8 detector array has been realized by growing large-area PtSe$_2$ combined with Si; a detectivity of 1.26×10^{13} Jones was achieved in the near-infrared band at 810 nm.[75]

2D materials are excellent candidates for curved image sensors, because their layered properties permit them to bend. 2D material detectors on flexible substrates can cover a broad spectral range from UV to IR.[76] A hemispherical structure can be assembled by bending a flexible substrate.[66] Therefore, many research studies have been conducted on 2D flexible detectors.[76] Choi et al. applied a MoS$_2$/graphene heterojunction photodetector[66] to an artificial retina in an aberration-free single-lens imaging system, as shown in Figure 6G. A flexible photodetector based on a graphene/MoS$_2$ vdW heterojunction achieved a photo-responsivity of 45.5 A W^{-1}.[77] The photodetector array and ultrathin neural-interfacing electrodes were integrated on a flexible printed circuit board to form a soft-implant optoelectronic device, as shown in Figure 6H. Live experiments on rats verified the effect of the device on the retina. The flexible device successfully responded to external light pulses and detected both spikes in the rat visual cortex and local field potential changes on the spikes. This was the first time that a bionic eye based on 2D materials was realized. Table 2 summarizes the scale and performance of various photodetector arrays based on 2D materials.

Moreover, black phosphorus[80] (bP), a narrow-bandgap semiconductor material, is widely used in room-temperature mid-infrared detection because of its high mobility and 0.3 eV bandgap.[10] Although CVD can be used to synthesize bP from bottom to top, only a few layers of bP thin films with a transverse size of tens of microns can be obtained.[81,82] Wu et al. reported a controlled PLD method that can directly synthesize at the centimeter scale a few layers with high crystallinity and high homogeneity.[83] Molecular dynamics (MD) simulation results show that the pulsed laser, unlike conventional thermal-assisted evaporation, promotes the uniform distribution of bP clusters in the physical vapor, thus reducing the formation energy of the bP phase and realizing the large-scale growth of bP thin films with fewer layers. A cm-level field-effect transistor (FET) array based on a few-layer bP film was fabricated. It exhibited good electrical characteristics in terms of carrier mobility and current switch ratio: the carrier mobility reached 213 cm^2 V^{-1} s^{-1} at 295 K and 617 cm^2 V^{-1} s^{-1} at 250 K, performance comparable to that previously reported for mechanically stripped bP. However, a bP photodetector array has not yet been reported. The optoelectronic properties of large-scale bP films cannot be compared with those of bP films obtained from mechanical exfoliation.

4.2 | Accurate characterization of quality factor of 2D-material photodetector

Photodetectors are classified as photoconductive or photovoltaic, depending on their structure. The photoconductive type is easy to manufacture and requires a current drive, but its dark current is higher; the photovoltaic type can work under zero bias conditions and has low noise, but its manufacturing process is more complicated. The performance quality factors of the two types of photodetectors are also different. Therefore, it is important to standardize the evaluation methods. Although the photoelectric properties of 2D materials are almost always reported to be excellent, these test data are frequently obtained through non-standard measurement methods, often using non-irradiation light sources such as lasers. The active area of the device is often also not clearly defined. Nonstandard measurement methods have seriously hindered the development of photodetectors, making it impossible to compare measured performances with those of commercial devices. Therefore, an accurate characterization method for the quality factor of a 2D photodetector is included in this paper.

2D material photodetectors often perform poorly in low-light detection. Most devices fail to detect photo signals in standard measurements using blackbody light sources. The responsivity and detectivity of some devices will drop by two or three orders of magnitude in standard measurements. Taking the specific detectivity of an important index of the infrared detector as an example,

Materials	Scale	Wavelength (nm)	R (A/W)	D^* (Jones)	Refs.
MoS_2	4×4	405	0.5	—	77
MoS_2	5×5	637	24	—	26
MoS_2	32×32	532	150	10^{10}	64
SnSe	≈ 1000	980	<1	10^9	78
MoS_2/graphene	1500	515	<1	—	66
Se_xTe_{1-x}	42×42	1550	1.5	10^{10}	65
$PdSe_2$/Si	4×4	780	<1	10^{13}	70
$PtSe_2$/Si	8×8	810	—	10^{13}	74
PdS/graphene	388×288	NIR	—	10^{12}	25
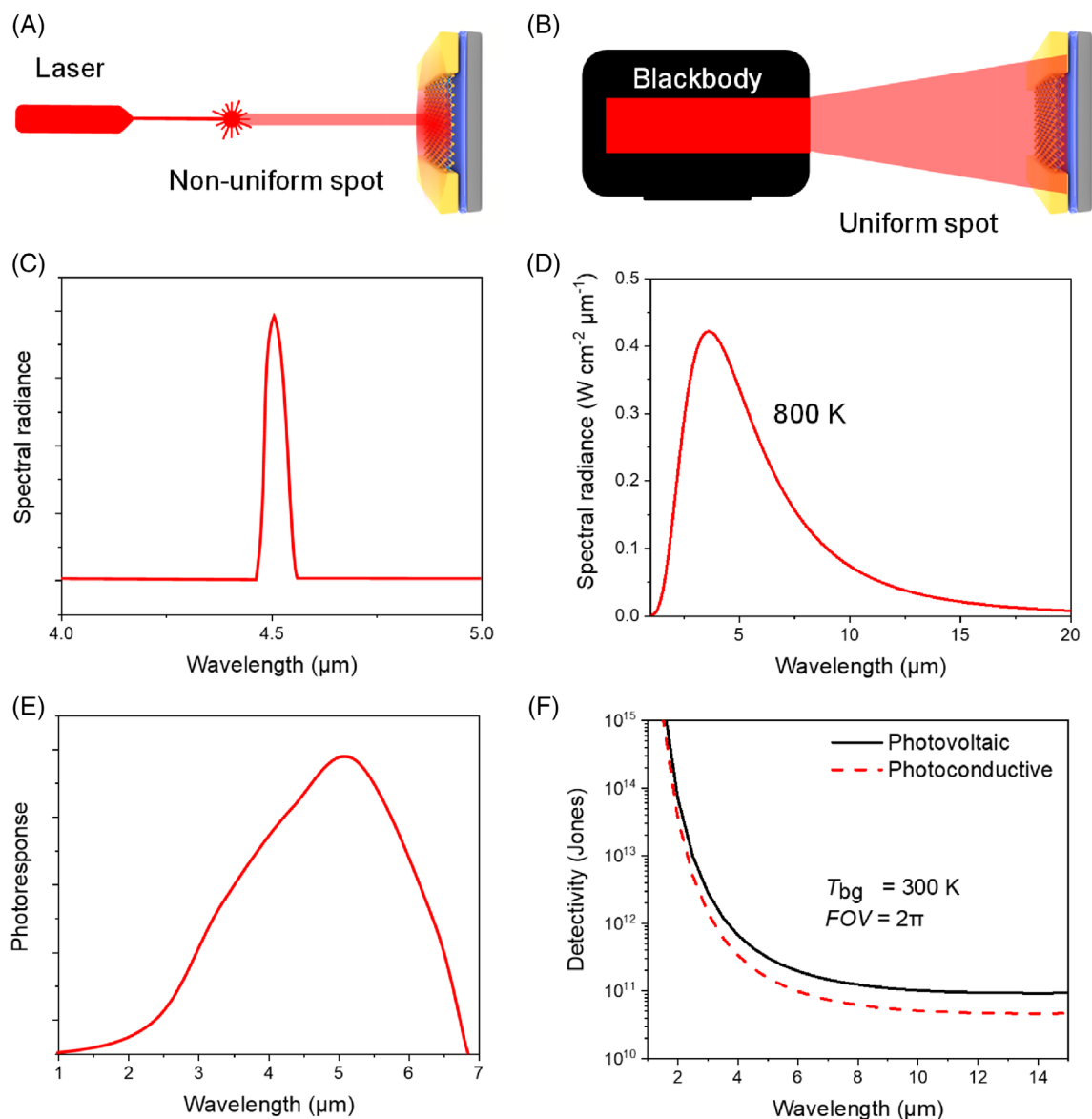
F I G U R E 7 Accurate characterization of quality factor of two-dimensional material photodetector. (A) Schematic diagram of photoresponse of laser test. (B) Schematic diagram of photoresponse of blackbody test. (C) Schematic diagram of the spectral distribution of a monochromatic laser. (D) Schematic diagram of 800 K blackbody spectral distribution. (E) Photodetector spectral response calculated using g-factor. (F) The detectivity limit of the photodetector limited by background radiation at 300 K (BLIP) when the field of view is 2π

the specific detectivity of the commercial thermistor is 10^8 Jones. Meaningful research work should be higher than the detectivity of the thermal detector. Therefore, the calculation of noise should also be standardized. The dark current should be carefully used to calculate the noise equivalent power (NEP).

One of the important indicators of the photodetector is the responsivity (R),[84] which is the ratio of the output photocurrent I_{ph} to the input optical power P:

$$R = \frac{I_{ph}}{P}. \tag{1}$$

At present, this quantity is usually measured by either the monochrome laser test or the mixed-light blackbody test.[85] Although the monochromatic property of the laser test is good (Figure 7C), the spot has a Gaussian distribution in the infrared stage and its size cannot be accurately measured, resulting in a calculation error of several orders of magnitude (Figure 7A). We believe that the uniform blackbody test should be the main test in the infrared band (Figure 7B). The device is located at a distance L from the aperture that can vary with the slide. The total incident power on the device surface can be calculated using the formula

$$P = \frac{\alpha\varepsilon\sigma\left(T^4 - T_0^4\right)A}{\pi L^4} A_n \tag{2}$$

where α is the modulation factor, ε is the average emissivity of the blackbody radiation source ($\varepsilon = 0.9$ for our source), and σ is the Stefan–Boltzmann constant. T is the temperature of the blackbody radiation source, and T_0 is the room temperature (300 K). A is the area of the blackbody radiation source, and A_n is the device area. The photodetector absorbs the incident photons to generate a voltage signal, which is converted by the current preamplifier to obtain the final current signal I. The current response rate R of the device is I/P. To obtain more accurate current signals, a chopper is introduced into the test optical-path diagram to modulate the radiation spectrum, and a phase-locked amplifier reads the signal to filter system noise and background noise. From this, we can obtain the blackbody responsivity.

To obtain the relationship between the responsivity of the detector and the radiation wavelength, a spectral response test is required. In this study, grating and Fourier transform infrared (FTIR) spectroscopy were combined for testing. In general, the blackbody response test obtains the blackbody responsivity R_b and blackbody detectivity D^* (blackbody), and the FTIR test obtains the relative responsivity $R'(\lambda)$. The ratio of the blackbody

responsivity and the peak responsivity of an infrared detector is constant, which is the g factor. Thus, by calculating the factor g, we can obtain the peak responsivity $R(\lambda_p)$ and peak detectivity $D^*(\lambda_p)$.[86] To calculate g, we first note that by definition

$$R'(\lambda) = \frac{R(\lambda)}{R(\lambda_p)}. \tag{3}$$

Because blackbody radiation has a continuous spectrum and the emissivity of each wavelength is different (Figure 7D), the signal produced by the photodetector is the sum of the signals produced by each wavelength of radiation:

$$R_b = \frac{\int_0^\infty \phi(\lambda)R(\lambda)d\lambda}{\int_0^\infty \phi(\lambda)d\lambda} = R(\lambda_p)\frac{\int_0^\infty \phi(\lambda)R(\lambda)/R(\lambda_p)d\lambda}{\int_0^\infty \phi(\lambda)d\lambda}$$
$$= R(\lambda_p)\frac{\int_0^\infty \phi(\lambda)R'(\lambda)d\lambda}{\int_0^\infty \phi(\lambda)d\lambda}, \tag{4}$$

where $\phi(\lambda)$ is the blackbody radiation power distribution. Then the g factor is

$$g = \frac{R_b}{R(\lambda_p)} \frac{\int_0^\infty R'(\lambda)\phi(\lambda)d\lambda}{\int_0^\infty \phi(\lambda)d\lambda} \tag{5}$$

The response spectrum obtained by the Fourier spectrometer is only a relative response spectrum, and the response rate of the current spectrum of the photodetector can be obtained after blackbody-response calibration. The relationship between the peak current photoresponse $R(\lambda_p)$ and the blackbody photoresponse R_b of the photodetector is as follows (Figure 7E):

$$R(\lambda_p) = \frac{R_b}{g}. \tag{6}$$

The normalized detectivity (D^*) is defined as the detector output signal-to-noise ratio (SNR) generated by unit incident power when normalized to a unit photosensitive surface and unit noise equivalent power (NEP); its unit is the Jones (i.e., cm $Hz^{1/2}$ W^{-1}).[87] The best performance can be obtained when the detector and amplifier noise is low. Under these conditions, the main noise source of the detector is the discrete radiation field in the detection process, and the main noise is photon noise. This limiting value of detector performance is called the background limit. From the particle noise formula, the BLIP (background limited infrared photodetector) D^* of the photovoltaic and photoconductive types at 300 K can be deduced,[87] as shown in Figure 7F. In principle, the

Roadmap for next-generation machine vision systems based on 2D Material photodetectors

System

Challenges
- High density interconnection technology
- System integrity

Potential solutions
- Using scale advantages of low-dimensional materials to achieve high-density interconnection
- Sense-memory-computational integration

Device

Challenges
- Performance evaluation method is not unified
- Photonic memristor has not been studied deeply enough

Potential solutions
- Standard device performance evaluation criteria

Material

Challenges
- Controlled high quality large area material growth technology
- Large area material transfer process

Potential solutions
- Application of MBE, MOCVD and other high quality epitaxial technology
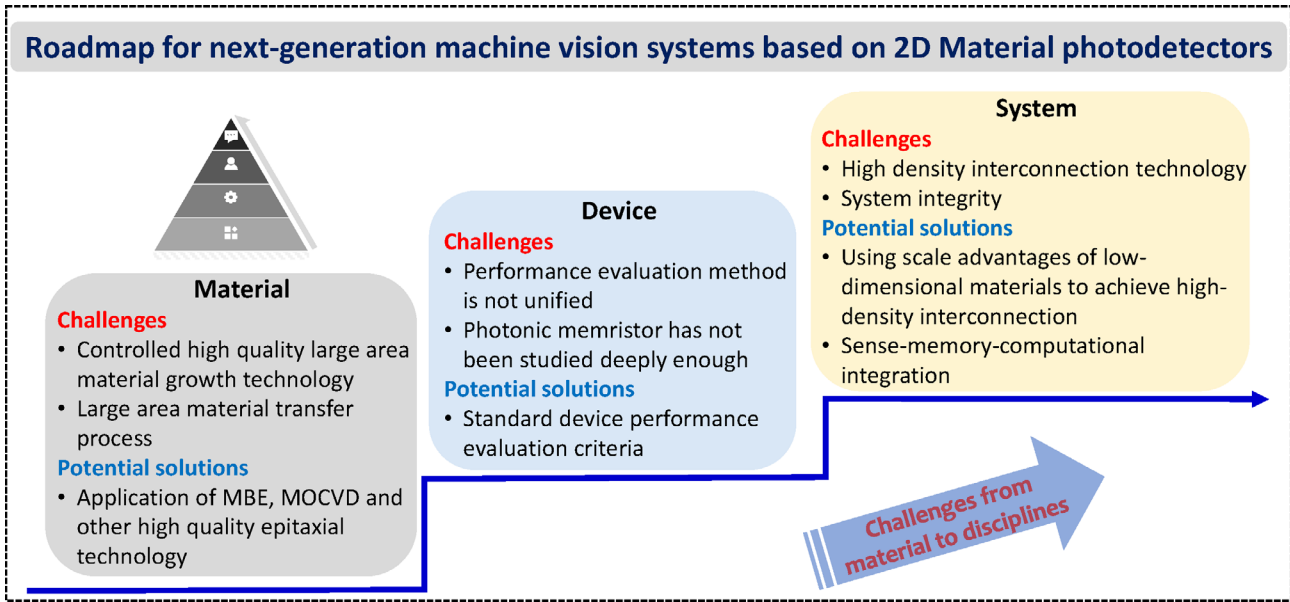
Challenges from material to disciplines

FIGURE 8 Roadmap for MVS based on 2D material photodetectors

BLIP is the performance limit of the detectors (Figure 7F).

We also calculated the D^* of some low-dimensional material detectors for comparison. D^* is usually expressed as follows:

$$D^* = \frac{\sqrt{A\Delta f}}{\text{NEP}} = R \frac{(A\Delta f)^{1/2}}{i_n}. \qquad (7)$$

Here, i_n is the noise current, which is different for photoconductive and photovoltaic detectors, Δf is the integration time (1 s). The gain of photosensitive diodes based on 2D materials is usually higher than 1. In the dark state of the photoconductive detector, the two main noises are thermal noise (i_t) and composite noise (i_{gr}). Both of these noises are related to the internal gain, so it is not accurate to calculate D^* using the dark current I_d for an optical detector with gain; it should be divided by the gain. Failure to divide by the gain is why reported values of low-dimensional-material photoconductive-detector performance are in general artificially high. Of course, the response time of high-gain photodetectors will be affected by gain to different degrees, depending on the source of the gain. For example, the gain of a high-gain avalanche photodetector comes from avalanche breakdown, so it has little effect on the response time. By contrast, gain caused by an internal trap state has a considerable impact on the response time.

The field of photodetectors based on 2D materials is booming and has vast potential.[84] However, accurate test methods are urgently needed to compare the performance indicators of various devices. The common practice of using dark current and laser tests to define device detection rates is inadequate, because it ignores key parameters, such as spot size, Gaussian distribution of optical power, and the gain relationship between dark current and noise in the gain device. These omissions may overestimate D^* by orders of magnitude. Therefore, we suggest that the SNR and NEP of the device can be directly characterized by the combined measurement of the noise spectral density and irradiated photocurrent at a fixed modulation frequency (and 1 Hz bandwidth). Such a strict calculation process is essential for verifying high D^* values and promoting the application of 2D photodetectors.

5 | CONCLUSION

2D material photodetectors provide solutions for next-generation machine vision problems. However, 2D materials are not yet feasible for practical use in intelligent photoelectric sensors. Technical issues that still need to be addressed include techniques for the stable growth of large-scale crystals, standardization of evaluation methods, and large-scale heterogeneous integration techniques. Existing intelligent sensing based on 2D materials has a relatively low scale, and system completeness is the main issue for intelligent sensing. A roadmap for next-generation MVSs based on 2D materials is shown in Figure 8. Theoretically, the three-dimensional integration of low-dimensional and other materials is possible; this would significantly improve system area efficiency, but

achieving high-density heterogeneous integration will require advances in selective 2D-material etching. In addition, the connection density of the alignment circuit should be considered. In terms of algorithms, existing MVSs based on 2D materials mainly focus on the improvement of recognition accuracy. However, accuracy is difficult to compare between systems with different structures, and a new quality factor needs to be introduced to evaluate next-generation MVSs based on 2D materials. In terms of hardware, the trend has been toward reduction rather than elimination of the AD module and auxiliary modules. For example, in Mennel's research,[9] external memory is required to store the weight value, and a computer is required to calculate the loss and classification functions. In the future, cutting-edge machine learning algorithms and neural network structures will be integrated into image sensors. Therefore, the development of new MVS architecture requires cross-domain exchanges between researchers and the collaboration of multiple disciplines, including materials science, semiconductor physics, computer science, and data science.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Zhiyong Fan https://orcid.org/0000-0002-5397-0129
Peng Zhou https://orcid.org/0000-0002-7301-1013
Weida Hu https://orcid.org/0000-0001-5278-8969

REFERENCES

1. Mead C. Neuromorphic electronic systems. *Proc IEEE*. 1990;78: 1629-1636.
2. Kyuma K, Lange E, Ohta J, Hermanns A, Banish B, Oita M. Artificial retinas — fast, versatile image processors. *Nature*. 1994;372:197-198.
3. Yang S-H, Kim K-B, Kim E-J, Baek K-H, Kim S. An ultra low power CMOS motion detector. *IEEE Trans. Broadcast Telev. Receivers*. 2009;55(4):2425-2430.
4. Wang S, Wang C-Y, Wang P, et al. Networking retinomorphic sensor with memristive crossbar for brain-inspired visual perception. *Natl Sci Rev*. 2021;8(2):nwaa172.
5. Norton P, Campbell J, Horn S, Reago D. Third-generation infrared imagers. In: Andresen BF, Fulop GF, Strojnik M, eds. *Infrared Technology and Applications XXVI*. Vol 4130. Proceedings of the Society of Photo-Optical Instrumentation Engineers (Spie). International Society for Optics and Photonics; 2000:226-236.
6. Bullock J, Amani M, Cho J, et al. Polarization-resolved black phosphorus/molybdenum disulfide mid-wave infrared photodiodes with high detectivity at room temperature. *Nat. Photonics*. 2018;12:601-607.
7. Ye L, Li H, Chen Z, Xu J. Near-infrared photodetector based on MoS_2/black phosphorus heterojunction. *ACS Photonics*. 2016;3: 692-699.
8. Guo Q, Pospischil A, Bhuiyan M, et al. Black phosphorus mid-infrared photodetectors with high gain. *Nano Lett*. 2016;16: 4648-4655.
9. Mennel L, Symonowicz J, Wachter S, Polyushkin DK, Molina-Mendoza AJ, Mueller T. Ultrafast machine vision with 2D material neural network image sensors. *Nature*. 2020;579: 62-66.
10. Chen Y, Wang Y, Wang Z, et al. Unipolar barrier photodetectors based on van der Waals heterostructures. *Nat. Electron*. 2021;4:357-363.
11. Liu Y, Huang Y, Duan X. Van der Waals integration before and beyond two-dimensional materials. *Nature*. 2019;567: 323-333.
12. Chhowalla M, Jena D, Zhang H. Two-dimensional semiconductors for transistors. *Nat. Rev. Mater*. 2016;1:15.
13. Koppens FHL, Mueller T, Avouris P, Ferrari AC, Vitiello MS, Polini M. Photodetectors based on graphene, other two-dimensional materials and hybrid systems. *Nat Nanotechnol*. 2014;9:780-793.
14. Liu Y, Weiss NO, Duan X, Cheng H-C, Huang Y, Duan X. Van der Waals heterostructures and devices. *Nat Rev Mater*. 2016;1:16042.
15. Behura SK, Wang C, Wen Y, Berry V. Graphene–semiconductor heterojunction sheds light on emerging photovoltaics. *Nat Photonics*. 2019;13:312-318.
16. Akinwande D, Huyghebaert C, Wang C-H, et al. Graphene and two-dimensional materials for silicon technology. *Nature*. 2019; 573:507-518.
17. Amer M, Maul T. A review of modularization techniques in artificial neural networks. *Artif Intell Rev*. 2019;52:527.
18. Chai Y. In-sensor computing for machine vision. *Nature*. 2020; 579:32-33.
19. Liu C, Chen H, Wang S, et al. Two-dimensional materials for next-generation computing technologies. *Nat Nanotechnol*. 2020;15:545-557.
20. Lichtsteiner P, Posch C, Delbruck T. A 128 × 128 120 dB 15 μs latency asynchronous temporal contrast vision sensor. *IEEE J. Solid-State Circuits*. 2008;43:566-576.
21. Brady TF, Konkle T, Alvarez GA, Oliva A. Visual long-term memory has a massive storage capacity for object details. *Proc Natl Acad Sci U S A*. 2008;105:14325-14329.
22. Euler T, Haverkamp S, Schubert T, Baden T. Retinal bipolar cells: elementary building blocks of vision. *Nat Rev Neurosci*. 2014;15:507-519.
23. Pospischil A, Furchi MM, Mueller T. Solar-energy conversion and light emission in an atomic monolayer p–n diode. *Nat Nanotechnol*. 2014;9:257-261.
24. Baugher BWH, Churchill HOH, Yang YF, Jarillo-Herrero P. Optoelectronic devices based on electrically tunable p-n diodes

in a monolayer dichalcogenide. *Nat Nanotechnol.* 2014;9: 262-267.

25. Goossens S, Navickaite G, Monasterio C, et al. Broadband image sensor array based on graphene–CMOS integration. *Nat. Photonics.* 2017;11:366-371.

26. Yang C-C, Chiu K-C, Chou C-T, et al. Enabling monolithic 3D image sensor using large-area monolayer transition metal dichalcogenide and logic/memory hybrid 3D+IC. Paper presented at: 2016 IEEE Symposium on VLSI Technology; June 14-16, 2016; Honolulu, HI:IEEE.

27. Wang L, Meric I, Huang PY, et al. One-dimensional electrical contact to a two-dimensional material. *Science.* 2013;342:614-617.

28. Guo H-W, Hu Z, Liu Z-B, Tian J-G. Stacking of 2D materials. *Adv Funct Mater.* 2021;31:2007810.

29. Lo SCB, Chan HP, Lin JS, Li H, Freedman MT, Mun SK. Artificial convolution neural network for medical image pattern recognition. *Neural Netw.* 1995;8:1201-1214.

30. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015; 521:436-444.

31. Biswas A, Chandrakasan AP. CONV-SRAM: An energy-efficient SRAM with in-memory dot-product computation for low-power convolutional neural networks. *IEEE J Solid-State Circuits.* 2019;54:217-230.

32. Kim NS, Chen D, Xiong J, Hwu W-MW. Heterogeneous computing meets near-memory acceleration and high-level synthesis in the post-moore era. *IEEE Micro.* 2017;37:10-18.

33. Lhostis S, Farcy A, Deloffre E, et al. Reliable 300 mm wafer level hybrid bonding for 3D stacked CMOS image sensors. *2016 IEEE 66th Electronic Components and Technology Conference (ECTC).* IEEE; 2016:869-876.

34. Kagawa Y, Iwamoto H, 3D Integration Technologies for the Stacked CMOS Image Sensors. Paper presented at: 2019 International 3D Systems Integration Conference (3DIC); October 8-10, 2019; Sendai, Japan.

35. Goto M, Hagiwara K, Iguchi Y, et al. Three-dimensional integrated circuits and stacked CMOS image sensors using direct bonding of SOI layers. Paper presented at: 2015 International 3D Systems Integration Conference (3DIC); August 31-September 2, 2015; Sendai, Japan.

36. Tsugawa H, Takahashi H, Nakamura R, et al. Pixel/DRAM/logic 3-layer stacked CMOS image sensor technology. Paper presented at: 2017 IEEE International Electron Devices Meeting (IEDM); December 2-6, 2017; San Francisco, CA.

37. Floreano D, Pericet-Camara R, Viollet S, et al. Miniature curved artificial compound eyes. *Proc Natl Acad Sci U S A.* 2013;110:9267-9272.

38. Song YM, Xie Y, Malyarchuk V, et al. Digital cameras with designs inspired by the arthropod eye. *Nature.* 2013;497:95-99.

39. Jayachandran D, Oberoi A, Sebastian A, et al. A low-power biomimetic collision detector based on an in-memory molybdenum disulfide photodetector. *Nat Electron.* 2020;3:646-655.

40. Wang C-Y, Liang S-J, Wang S, et al. Gate-tunable van der Waals heterostructure for reconfigurable neural network vision sensor. *Sci Adv.* 2020;6(26):aba6173.

41. Dowling JE, Boycott BB. Organization of the primate retina: electron microscopy. *Proc R Soc Lond B Biol Sci.* 1966;166:80-111.

42. Demb JB. Cellular mechanisms for direction selectivity in the retina. *Neuron.* 2007;55:179-186.

43. Masland RH. The fundamental plan of the retina. *Nat Neurosci.* 2001;4:877-886.

44. Tong L, Peng Z, Lin R, et al. 2D materials–based homogeneous transistor-memory architecture for neuromorphic hardware. *Science.* 2021;373(6561):1353-1358.

45. Jeong KH, Kim J, Lee LP. Biologically inspired artificial compound eyes. *Science.* 2006;312:557-561.

46. Zhang L, Pasthukova N, Yao Y, et al. Self-suspended nanomesh scaffold for ultrafast flexible photodetectors based on organic semiconducting crystals. *Adv Mater.* 2018;30(28):201801181.

47. Lee W, Lee J, Yun H, et al. High-resolution spin-on-patterning of perovskite thin films for a multiplexed image sensor array. *Adv Mater.* 2017;29:1702902.

48. Zhou F, Zhou Z, Chen J, et al. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat Nanotechnol.* 2019;14:776-782.

49. Gong F, Luo W, Wang J, et al. High-sensitivity floating-gate phototransistors based on WS_2 and MoS_2. *Adv Funct Mater.* 2016;26:6084-6090.

50. Du C, Cai F, Zidan MA, Ma W, Lee SH, Lu WD. Reservoir computing using dynamic memristors for temporal information processing. *Nat Commun.* 2017;8:2204.

51. Zhou F, Liu Y, Shen X, Wang M, Yuan F, Chai Y. Low-voltage, optoelectronic $CH_3NH_3PbI_{3-x}Cl_x$ memory with integrated sensing and logic operations. *Adv Funct Mater.* 2018;28(15): 201800080.

52. Wang P, Xia H, Li Q, et al. Sensing infrared photons at room temperature: From bulk materials to atomic layers. *Small.* 2019;15(46):201904396.

53. Liu C, Wang L, Qi J, Liu K. Designed growth of large-size 2D single crystals. *Adv Mater.* 2020;32:2000046.

54. He T, Wang Z, Cao R, et al. Extrinsic Photoconduction Induced Short-Wavelength Infrared Photodetectors Based on Ge-Based Chalcogenides. *Small.* 2021;17(4):202006765.

55. Jiang J, Ling C, Xu T, et al. Defect Engineering for Modulating the Trap States in 2D Photoconductors. *Adv Mater.* 2018;30:e1804332.

56. Wang H, Wang G, Bao P, et al. Controllable synthesis of submillimeter single-crystal monolayer graphene domains on copper foils by suppressing nucleation. *J Am Chem Soc.* 2012;134:3627-3630.

57. Yan Z, Lin J, Peng Z, et al. Toward the synthesis of wafer-scale single-crystal graphene on copper foils. *ACS Nano.* 2013;7(3): 2872.

58. Zhou H, Yu WJ, Liu L, et al. Chemical vapour deposition growth of large single crystals of monolayer and bilayer graphene. *Nat Commun.* 2013;4:2096.

59. Hao Y, Bharathi MS, Wang L, et al. The role of surface oxygen in the growth of large single-crystal graphene on copper. *Science.* 2013;342:720-723.

60. Gan L, Luo Z. Turning off hydrogen to realize seeded growth of subcentimeter single-crystal graphene grains on copper. *ACS Nano.* 2013;7:9480-9488.

61. Guo W, Jing F, Xiao J, Zhou C, Lin Y, Wang S. Oxidative-etching-assisted synthesis of centimeter-sized single-crystalline graphene. *Adv Mater.* 2016;28:3152-3158.

62. Vlassiouk IV, Stehle Y, Pudasaini PR, et al. Evolutionary selection growth of two-dimensional materials on polycrystalline substrates. *Nat Mater.* 2018;17:318-322.

63. Konstantatos G, Badioli M, Gaudreau L, et al. Hybrid graphene-quantum dot phototransistors with ultrahigh gain. *Nat Nanotechnol.* 2012;7:363-368.

64. Jang H, Liu C, Hinton H, et al. An atomically thin optoelectronic machine vision processor. *Adv Mater.* 2020;32:2002431.

65. Tan C, Amani M, Zhao C, et al. Evaporated Se$_x$Te$_{1-x}$ thin films with tunable bandgaps for short-wave infrared photodetectors. *Adv Mater*. 2020;32:2001329.

66. Choi C, Choi MK, Liu S, et al. Human eye-inspired soft optoelectronic device using high-density MoS$_2$-graphene curved image sensor array. *Nat Commun*. 2017;8:1664.

67. Wang Q, Li N, Tang J, et al. Wafer-scale highly oriented monolayer MoS$_2$ with large domain sizes. *Nano Lett*. 2020;20: 7193-7199.

68. Hoang Danh P, Kim Y, Lee J, et al. Ultraclean and direct transfer of a wafer-scale MoS$_2$ thin film onto a plastic substrate. *Adv Mater*. 2017;29(7):201603928.

69. Lin H, Zhu Q, Shu D, et al. Growth of environmentally stable transition metal selenide films. *Nat Mater*. 2019;18:602-607.

70. Zeng L-H, Wu D, Lin S-H, et al. Controlled synthesis of 2D palladium diselenide for sensitive photodetector applications. *Adv Funct Mater*. 2019;29:1806878.

71. Serna MI, Yoo SH, Moreno S, et al. Large-area deposition of MoS$_2$ by pulsed laser deposition with in situ thickness control. *ACS Nano*. 2016;10:6054-6061.

72. Tan LK, Liu B, Teng JH, et al. Atomic layer deposition of a MoS$_2$ film. *Nanoscale*. 2014;6:14002-14002.

73. Kang K, Xie S, Huang L, et al. High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity. *Nature*. 2015;520:656-660.

74. Li J, Wang S, Jiang Q, et al. Single-crystal MoS$_2$ monolayer wafer grown on Au (111) film substrates. *Small*. 2021;6(18): 10584-10588.

75. Ma M, Chen H, Zhou K, et al. Multilayered PtSe$_2$/pyramid-Si heterostructure array with light confinement effect for high-performance photodetection, image sensing and light trajectory tracking applications. *J Mater Chem C*. 2021;9:2823-2832.

76. Kang P, Wang MC, Knapp PM, Nam S. Crumpled graphene photodetector with enhanced, strain-tunable, and wavelength-selective photoresponsivity. *Adv Mater*. 2016;28:4639-4645.

77. De Fazio D, Goykhman I, Yoon D, et al. High responsivity, large-area graphene/MoS$_2$ flexible photodetectors. *ACS Nano*. 2016;10:8252-8262.

78. Park H, Liu N, Kim BH, et al. Exceptionally uniform and scalable multilayer MoS$_2$ phototransistor array based on large-scale MoS$_2$ grown by RF sputtering, electron beam irradiation, and sulfurization. *ACS Appl Mater Interfaces*. 2020;12: 20645-20652.

79. Hao L, Du Y, Wang Z, et al. Wafer-size growth of 2D layered SnSe films for UV-Visible-NIR photodetector arrays with high responsivity. *Nanoscale*. 2020;12:7358-7365.

80. Li L, Yu Y, Ye GJ, et al. Black phosphorus field-effect transistors. *Nat Nanotechnol*. 2014;9:372-377.

81. Smith JB, Hagaman D, Ji H-F. Growth of 2D black phosphorus film from chemical vapor deposition. *Nanotechnology*. 2016;27: 215602.

82. Xu Y, Shi X, Zhang Y, et al. Epitaxial nucleation and lateral growth of high-crystalline black phosphorus films on silicon. *Nat Commun*. 2020;11:1330.

83. Wu Z, Lyu Y, Zhang Y, et al. Large-scale growth of few-layer two-dimensional black phosphorus. *Nat Mater*. 2021;1-7.

84. Fang Y, Armin A, Meredith P, Huang J. Accurate characterization of next-generation thin-film photodetectors. *Nat. Photon*. 2019;13:1-4.

85. Peng M, Xie R, Wang Z, et al. Blackbody-sensitive room-temperature infrared photodetectors based on low-dimensional tellurium grown by chemical vapor deposition. *Sci Adv*. 2021;7 (16):abf7358.

86. Vincent JD, Hodges S, Vampola J, Stegall M, Pierce G. *Fundamentals of Infrared and Visible Detector Operation and Testing*. John Wiley & Sons; 2015.

87. Rogalski A. *Infrared Detectors*. CRC press; 2000.

AUTHOR BIOGRAPHIES

Peisong Wu obtained his Bachelor degree at Jilin University in China in 2017. He is currently a Ph.D. student in National Laboratory for Infrared Physics at Shanghai Institute of Technical Physics Chinese Academy of Sciences under the supervision of Prof. Weida Hu. His research interests are fabrication and application of smart infrared photodetectors based on low-dimensional materials.

Peng Zhou is a full professor at the state key laboratory of ASIC and system, school of microelectronics, Fudan University, China. He received his B.S. (2000) and Ph.D. (2005) degrees in Physics from Fudan University, China. Currently, Prof. Zhou is interested in novel high-efficiency and low-power electronic devices based on 2D layered materials, focusing on the application of 2D materials in memory, including 2D flash memory, floating-gate memory, semi-floating-gate quasi-nonvolatile memory, synaptic electronics, and neuromorphic systems.

Weida Hu received his Ph.D. with honors in microelectronics and solid-state electronics from the Shanghai Institute of Technology Physics (SITP), Chinese Academy of Sciences (CAS), in 2007. He is currently a full professor on fabrication and characterization of IR photodetectors and their smart chips in SITP, CAS.