Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ambipolar photoresponsive infrared detector arrays for energy-efficient in-sensor optoelectronic computing

Zhang Zhiyong

zyzhang@pku.edu.cn

Peking University https://orcid.org/0000-0003-1622-3447

Qianxi Yang

Key Laboratory for the Physics and Chemistry of Nanodevices and Center for Carbon-based Electronics, School of Electronics, Peking University, Beijing, China

Hongkun Duan

School of Integrated Circuits, Beijing University of Posts and Telecommunications, Beijing, China. Key Laboratory of Luminescence & Optical Information, Ministry of Education, School of Physical Sci

XiaoLu Xia

Key Laboratory of Luminescence & Optical Information, Ministry of Education, School of Physical Science and Engineering, Beijing Jiaotong University, Beijing, China

Shaoyuan Zhou

Key Laboratory for the Physics and Chemistry of Nanodevices and Center for Carbon-based Electronics, School of Electronics, Frontiers Science Center for Nano-optoelectronics, Peking University, Beij

Zixuan Yang

Peking University

Tao Luo

Key Laboratory for the Physics and Chemistry of Nanodevices and Center for Carbon-based Electronics, School of Electronics, Peking University, Beijing, China

Yixiao Niu

School of Integrated Circuits, Huazhong University of Science and Technology, Hubei, China

Zhiyu Zhao

School of Integrated Circuits, Huazhong University of Science and Technology, Hubei, China

Yi Cao

Xiangtan university

yuyan Zhang

Peking University

Jianhua Jiang

Peking University

Jianbing Zhang

Huazhong University of Science & Technology https://orcid.org/0000-0003-0642-3939

Ying Wang

Key Laboratory of Luminescence & Optical Information, Ministry of Education, School of Physical Science and Engineering, Beijing Jiaotong University, Beijing, China. Chongqing Institute of Carbon-Ba

Yongfeng Wang

Research Center for Carbon-based Electronics and Key Laboratory for the Physics and Chemistry of Nanodevices, School of Electronics, Peking University, Beijing 100871 https://orcid.org/0000-0002-8171-3189

Article

Keywords: ambipolar photodetector, infrared kernel primitives, in-sensor computing, broadband photodetection, carbon nanotube

Posted Date: June 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9937908/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: There is **NO** Competing Interest.

Ambipolar photoresponsive infrared detector arrays for energy-efficient in-sensor optoelectronic computing

Qianxi Yang[1], Hongkun Duan[2], Xiaolu Xia[2], Shaoyuan Zhou[1], Zixuan Yang[1], Tao Luo[1], Yixiao Niu[3], Zhiyu Zhao[3], Yi Cao[1], Yuyan zhang[1], Jianhua Jiang[1], Jianbing Zhang[3], Ying Wang[2,*], Yongfeng Wang[1,*] and Zhiyong Zhang[1,*]

[1]Key Laboratory for the Physics and Chemistry of Nanodevices and Center for Carbon-based Electronics, School of Electronics, Peking University, Beijing 100871, China.

[2]Key Laboratory of Luminescence & Optical Information Ministry of Education, School of Physical Science and Engineering, Beijing Jiaotong University, Beijing 100044, China.

[3]School of Integrated Circuits, Huazhong University of Science and Technology, Hubei 430074, China.

E-mail: yingw@bjtu.edu.cn; yongfengwang@pku.edu.cn; zyzhang@pku.edu.cn

Abstract

In-sensor optoelectronic computing offers a compelling route toward low-latency and energy-efficient machine vision, yet its expansion into the infrared spectrum remains fundamentally constrained by the absence of scalable and highly sensitive infrared kernel primitives (IKPs), which limits the application particularly under ultralow-light and complex conditions. Here we report highly efficient IKPs enabled by an ambipolar photoresponsive infrared photodetector consisting of an ambipolar field-effect transistor (AFET) and a heterojunction according to an optoelectrically decoupled architecture. Taking advantage of the symmetric ambipolar transport in Schottky-barrier (SB) carbon nanotube FET, the heterojunction-gated ambipolar FET (HGAFET) converts a directed photovoltage from the heterojunction gate into a positive photocurrent (PPC) or negative photocurrent (NPC) through adjusting the bottom gate voltage. The IKPs exhibit symmetric and sensitive ambipolar photoresponses, with both specific detectivities exceeding 10^{13} Jones, to broad spectrum light from visible-to-shortwave infrared (vis-SWIR), and then have been demonstrated in-sensor multiply-accumulate operations and analog-domain interframe difference computations. The proposed in-sensor computing system based on these IKPs enables efficient feature recognition and motion awareness, achieving recognition accuracies above 90% under low-light conditions while reducing inference latency and computational cost by factors of 16.2× and 73.0×, respectively, and can act as a hardware-level strategy for all-weather infrared edge vision.

Keywords: ambipolar photodetector, infrared kernel primitives, in-sensor computing, broadband photodetection, carbon nanotube

1 Introduction

Dynamic and static image detection spanning the visible-to-shortwave infrared (vis-SWIR) spectral range is essential for reliable target recognition and tracking in all-weather and challenging environments[1]. The existing imaging systems of this kind typically rely on multimodal sensors fusion[2], combining CMOS image sensors (CIS), infrared cameras[3], LiDAR[4], and high-dynamic-range (HDR) imaging[5]. Such approaches often oversample visual scenes, generating large volumes of redundant data that increase system latency and impose substantial downstream computational burdens. Event-driven imagers, inspired by biological vision, detect only temporal intensity changes, which contributes to significantly reducing data redundancy and achieving ultrahigh temporal resolution[6–8]. However, their pixel unit is inherently complex, commonly requiring multiple detectors and peripheral processing circuits, which makes pixel scaling difficult and fabrication more complex, which thereby severely constraining the spatial resolution and increasing the cost of the imaging chip.

Recent advances in emerging optoelectronic imaging sensors have enabled visual information to be captured and processed directly at the sensor level, supporting more compact and energy-efficient imaging systems[9,10]. Representatively, by exploiting gate-tunable photoresponse, photodetector arrays have been explored as hardware platforms for artificial neural networks (ANNs) and parallel computing, in which incident light is convolved with electrically encoded kernel primitives. Such in-sensor computing schemes have demonstrated static image pre-processing and recognition functions, including edge enhancement[11], spectral adaptation[12], data compression[13], image fusion[14], and noise reduction[15]. Motivated by motion perception, these paradigms have further evolved toward event- or frame-based architectures incorporating spiking neural networks (SNNs)[16,17], dynamic computing[18], visual adaptation[12,19], and neuromorphic synapses[20,21]. Nevertheless, all the reported kernel primitives failed to cover the SWIR band, restricting their applicability to all-weather, energy-efficient edge vision.

This limitation primarily stems from the challenge of fabricating a SWIR sensor with a high signal-to-noise ratio (SNR) and a stable, symmetric ambipolar

photoresponse simultaneously, thereby failing to provide effective infrared kernel primitives (IKPs) required for efficient in-sensor computing. Additionally, existing optoelectronic sensing units usually utilize complex structures such as split-gate[11,16,22,23] to achieve ambipolar photoresponse, which greatly increases the difficulty of pixel scaling down and is not easily compatible integration with silicon readout integrated circuits (ROICs). Therefore, a compact IKP, featuring a highly sensitive ambipolar response to vis-SWIR light, excellent fabrication compatibility and scaling down property, is urgently required to promote energy-efficient in-sensor optoelectronic computing for all-weather infrared edge vision.

In this work, we develop in-sensor IKPs using reconfigurable heterojunction-gated ambipolar field-effect transistor arrays based on carbon nanotube (CNT) films and semiconductor quantum dots according to an optoelectrically decoupled architecture. The fabricated single HGAFET can exhibit ambipolar photoresponse to ultra-low vis-SWIR light through adjusting back-gate voltage, as well as excellent fabrication compatibility to Si ROICs and scaling down property. These properties enable in-sensor multiply-accumulate (MAC) operations for static image processing, achieving recognition accuracies above 90%, particularly under low-light conditions. Moreover, the symmetric ambipolar output allows analog-domain interframe difference computation for dynamic perception, substantially simplifying pixel circuitry while reducing inference latency and computational cost. Combined with sensitive broadband detection and compatibility with monolithic three-dimensional (M3D) integration with Si ROICs (Fig. 1a), the demonstrated IKPs provide a scalable hardware foundation for infrared edge-vision systems requiring high spatial resolution and energy efficiency, with potential applications in unmanned aerial vehicles (UAVs), micro-robots, and satellite-based imaging.

2 Results and Discussion

2.1 Infrared Kernel Primitives Based on HGAFETs

Figure 1b schematically illustrates the implementation of IKPs for both static and

dynamic image processing using a HGAFET array. For compact and weight-constrained platforms such as UAVs, M3D integration (Fig. 1a) provides an effective route to reduce interconnect overhead while enabling high-throughput sensing–memory–computing stacks. Accordingly, HGAFET arrays intended for such M3D architectures must support wafer-scale fabrication, low-temperature processing compatible with Si back-end-of-line (BEOL) process, and operation with acceptable control complexity. To function as IKPs for in-sensor convolutional image processing, HGAFETs are required to exhibit a broadband ambipolar photoresponse, such that the output photocurrent can be expressed as the product of an optical input term and a gate-dependent weighting function. Crucially, this weighting function—defined through responsivity or other optoelectronic metrics—must vary deterministically with the applied gate voltage to enable reliable programmability and kernel encoding directly at the sensor level. Beyond static image processing, HGAFET-based IKPs should also support dynamic perception tasks, including analog-domain interframe differencing. Such operation relies on highly symmetric ambipolar photoresponses including positive photocurrent (PPC) and negative photocurrent (NPC) characteristics. As illustrated in Fig. 1b, each HGAFET pixel can be switched sequentially between NPC and PPC modes, and respectively generates $-I_{ph}$ and $+I_{ph}$ signals. Analog summation of these signals suppresses static background components while selectively enhancing motion-induced variations. A subsequent thresholding step produces a sparse motion mask that retains only motion-relevant information. This scheme enables concurrent extraction of static and dynamic visual cues with microsecond-scale latency and minimal circuit or algorithmic overhead (Supplementary Fig. 1).

Figure 1c shows the operating principle of an individual HGAFET, with the corresponding structural diagram presented in Fig. 2a. The device employs a p-i-n heterojunction to electrostatically gate a semiconductor channel that supports ambipolar charge transport (Methods and Supplementary Fig. 2) through a thin gate dielectric layer. The channel consists of a randomly oriented CNT thin film (Supplementary Figs. 3 and 4), which provides intrinsic ambipolarity with a moderate

bandgap (0.5–0.6 eV) and a mid-gap level of approximately 4.7 eV. Titanium (4.5 eV) is selected as source/drain contacts to form a mid-gap Schottky-barrier CNT FET, enabling similar injection efficiency for electrons and holes. In a local back-gate configuration, the CNT FET exhibits a characteristic V-shaped transfer curve ($I_{ds} - V_{bg}$) and near-linear output characteristics ($I_{ds} - V_{ds}$) at low bias (Supplementary Fig. 5), confirming the typical ambipolar transport with an on/off ratio exceeding 10^4.

Polarity-reconfigurable photoresponse is achieved by integrating a vertically stacked PbS quantum dot (QD)/ZnO heterostructure onto the top-gate dielectric (7 nm HfO_2), forming a photovoltaic gate for the ambipolar CNT FET. Under vis-SWIR illumination, photogenerated electrons accumulate at the ZnO/HfO_2 interface and generate a photovoltage to adjust carrier concentration of the CNT channel. Negative back-gate voltage V_{bg} tunes the Fermi level of CNT channel close to valence band and makes the SB for holes thinner, corresponding to p-type conduction, whereas positive V_{bg} tunes the Fermi level of CNT channel close to conduction band and makes the SB for electrons thinner, corresponding to n-type operation. As a result, photocurrents with different polarities are obtained even under the same optical excitation (inset of Fig. 1c). This optoelectrically decoupled mechanism suppresses optical–electrical crosstalk and relaxes the absorption-transport trade-off,[24] providing a suitable device framework for gate-programmable and highly sensitive IKPs in in-sensor optoelectronic computing.

2.2 Broadband Sensitive Ambipolar Photoresponse of HGAFET

The gate-tunable photoresponse of CNT-based HGAFETs was evaluated under infrared illumination, with 1300 nm used as a representative wavelength unless otherwise specified. The devices feature a channel area of 10 μm × 10 μm and a ZnO-defined gate length of 8 μm (Fig. 2a and Supplementary Fig. 6). Figure 2b shows the transfer characteristics measured under incident power densities (P_{light}) ranging from 0 to 37.4 mW cm^{-2}. In the p-type regime (V_{bg}<0), incident light induces a positive shift of the transfer curve (Fig. 1c, left) and yields a positive photocurrent (PPC; left panel in Fig. 2b), whereas in the n-type regime (V_{bg}>0), the same incident light results in a negative

photocurrent (NPC; right panel in Fig. 2b). This polarity switching originates from drain current change of the ambipolar FET in different modes (p/n-type branch) induced by the same photovoltage at the top gate (inset of Fig. 2b). Transient photoresponse measurements performed at V_{bg} = −0.5 V (p-type) and 0.1 V (n-type) under ultralow irradiances from 4.16 nW cm^{-2} to 1.92 μW cm^{-2} (Fig. 2c) confirm stable and reproducible ambipolar photoresponses. It is worth mentioning that the HGAFET exhibits obvious PPC/NPC responses (9.9 pA/-17.5 pA) to SWIR illumination with a power density as low as 4.16 nW cm^{-2}, reflecting the ultra-high sensitivity of the ambipolar photodetector with optoelectrically decoupled architecture. During the detecting or imaging process, the HGAFET is tuned into the subthreshold region by the applied V_{bg}, and then the photocurrent (see the detailed derivation process in the Methods) can be expressed as

$$I_{ph} = \left[\left(aP_{light} + 1\right)^{b/SS} - 1\right]\left(Ce^{\beta V_{bg}}\right) = I\left(P_{light}\right)W\left(V_{bg}\right) \quad (1)$$

where SS is the subthreshold swing of the FET, $a = \eta qA/(hvI_0)$, $b = 2.3\alpha k_B T/q$, η is the external quantum efficiency, A is the effective device area, I_0 is the reverse saturation current, α is the capacitive coupling coefficient, k_B is the Boltzmann constant, T is the temperature, hv is the photon energy, q is the elementary charge, C is a device-dependent constant, $\beta = q/(nk_B T)$, and n is the subthreshold slope factor (typically 1.2–2). Equation (1) can fit the experimentally measured $I_{ph}-P_{light}$ curves of the HGAFET in both PPC and NPC modes (Fig. 2d, e). The responsivity R, which is defined as $I_{ph}/(AP_{light})$, monotonically increases with the declining P_{light} in whatever PPC (inset in Fig. 2d) or NPC (inset in Fig. 2e) mode, which is consistent with the behavior of all reported photovoltaic detectors[25,26]. Furthermore, R increases with $|V_{bg}|$ at both PPC and NPC modes (insets in Fig. 2d and 2e), mainly owing to increasing transconductance (g_m) of top-gate CNT FETs induced by increased $|V_{bg}|$, which is confirmed through control measurements on dual-gated ambipolar CNT FETs in which the PbS QD/ZnO heterostructure is replaced by a metallic top gate (Supplementary Figs. 7 and 8). The monotonic dependence of R on $|V_{bg}|$ can be used

to construct back-gate voltage controlled weighting photodetection as in Fig. 2f, which shows photocurrent of the HGAFET at various V_{bg} varying from −0.9 V to 0.3 V under fixed illumination (1300 nm, 1.92 μW cm^{-2}). Mapping I_{ph} as a function of V_{bg} and P_{light} yields responsivity landscapes spanning more than five orders of magnitude (1−10^5 A W^{-1}) within ±2 V (Fig. 2g, h and Supplementary Fig. 9). Current noise spectra of the HGAFET were also measured at varying V_{bg} (Supplementary Fig. 10), and all show typical flicker (1/*f*) noise characteristics. We retrieve the V_{bg} dependent I_n curve at 10 Hz (Supplementary Fig. 10c), and the V-shaped dependence closely follows the transfer characteristics in Fig. 1c, indicating a strong correlation between current noise and dark current[24]. The resulting specific detectivity $D^* = R\sqrt{A\Delta f}/I_n$ also depends on V_{bg} (Fig. 3a, b), and the maximum value reaches 4.7 × 10^{13} Jones for NPC (V_{bg} = 0.4 V) and 2.6 × 10^{13} Jones for PPC (V_{bg} = −0.8 V), indicating that the HGAFET is among the most sensitive SWIR detectors in both operating polarities[26–35].

Owing to the broadband absorption of PbS QDs, HGAFETs maintain symmetric ambipolar photoresponse to light from 640 to 2100 nm (Fig. 3c and Supplementary Fig. 11). In contrast to trap-assisted[20,36] or photothermal polarity-switching mechanisms[37], the photocurrent polarity in HGAFETs is completely determined by V_{bg}, ensuring stable and reproducible operation across the vis-SWIR spectrum. As a result, the HGAFET exhibits almost symmetric $R - P_{light}$ curves between NPC and PPC modes at different wavelengths (650, 950, 1300, 1550, and 1700 nm; Supplementary Fig. 12) as shown in Fig. 3d. Benchmarking against representative ambipolar and unipolar photodetectors using two groups of key metrics—including detectable P_{light} *vs.* D^*, and responsivity *vs.* wavelength spectral coverage (Fig. 3e, f)[11,25–35,38–45]—highlights the performance advantages of the HGAFETs on sensitivity, bandwidth and ambipolarity, indicating the first ambipolar vis-SWIR detector with symmetry and high-sensitivity NPC and PPC responses, as well as excellent stability after two months of storage (Supplementary Fig. 13). Furthermore, rise and fall times of the HGAFETs are

measured and retrieved as 60–300 μs (Supplementary Fig. 13), which satisfy the requirements for in-sensor frame differencing and dynamic scenes involving tens-of-microseconds-level motion, such as obstacle detection in autonomous vehicles and fast industrial inspection. Further improvements are expected through adopting aligned CNT as channel, scaling down the channel length, and reducing defect densities in the QD and dielectric layers. The highly sensitive HGAFETs with excellent stability and sufficient speed are competent to act as building blocks for in-sensor computing.

2.3 Wafer-Scale HGAFET Arrays

The fabrication of HGAFETs includes preparation of CNT films and PbS QDs, fabrication of CNT FET and heterojunction involve low-temperature process (≤300 °C), and are fully compatible with the BEOL process of Si CMOS. To evaluate wafer-scale manufacturability, CNT-based HGAFET arrays were fabricated on a 4-inch Si/SiO_2 wafer, which contains 28 dies and 29,568 devices (Fig. 4a and Supplementary Fig. 14). The transfer curves of 64 randomly selected devices are shown in Fig. 4b, and exhibit threshold-voltage variations below 9% and on/off ratio variations below 4% across both p-type and n-type branches (Supplementary Fig. 15 and Supplementary Table 2-1), indicating good electrical reproducibility. Spatial maps of dark current and NPC/PPC photoresponse under 1300 nm illumination (Fig. 4c) further confirm the uniformity of ambipolar photoresponse across the wafer. Photocurrent distributions exhibit coefficients of variation ranging from 17% to 25% (Fig. 4d and Supplementary Table 2-2), exceeding those of purely electrical parameters. Such large variations are primarily attributed to thickness and packing-density fluctuations in the spin-coated PbS QD films, which are expected to be improved by optimizing the spin-coating process, for example by optimizing spin speed and solution concentration to improve the thickness uniformity, together with improved interfacial passivation to suppress local response variation. Therefore, the HGAFET platform enables wafer-scale fabrication with good uniformity potentially for array-level in-sensor computing.

2.4 Sensor-Level Static Image Processing

Presenting ambipolar photoresponse with high sensitivity and scalable fabrication, HGAFET arrays were employed as IKPs for programmable in-sensor convolution under low-light conditions. Through programming the weights by adjusting V_{bg} which should be connected to the Si ROICs, classical convolution operators, including Sobel, Prewitt, Laplacian, and mean filters, can be implemented at the pixel level within a 3 × 3 kernel (Supplementary Fig. 16). As shown in Fig. 5a, nine HGAFETs in each kernel are independently addressed via dedicated gate lines, while a peripheral digital-to-analog converter supplies multilevel V_{bg} values to assign positive, negative, or near-zero weights. Input images are partitioned into overlapping 3 × 3 patches and mapped onto the kernel pixels. Under illumination, the weighted photocurrents are summed along a shared drain line, producing convolution outputs directly in the analog domain and eliminating intermediate digitization. Continuous gate adjustment within ±2 V, combined with the wide ambipolar photocurrent responsivity of the HGAFETs (Fig. 2g, h), ensures sufficient dynamic range for accurate weight programming, establishing a direct physical mapping between device performance and convolution kernels.

Low-light performance was evaluated using a 478 × 359-pixel image from the Exclusively Dark Image Dataset. Owing to the ultrahigh sensitivity of the HGAFETs, reliable feature extraction is achieved at irradiance as low as 4.16 nW cm^{-2}, enabling clear identification of building contours, rail textures, and license plates under moonlight conditions (Fig. 5a and Supplementary Fig. 17). Conventional digital convolution pipelines, in contrast, exhibit strong illumination dependence, resulting in pronounced feature degradation under low brightness. Robustness against illumination variation was further examined using the augmented Yale Face Dataset across eight lighting conditions (Fig. 5b and Supplementary Fig. 18). When applying a Laplacian kernel, the HGAFET-based IKPs preserve high-fidelity edge information across all illumination levels, whereas conventional preprocessing degrades markedly at low brightness (Fig. 5c). Leveraging these illumination-stable features, a lightweight convolutional neural network (CNN; Fig. 5d) achieves 94.2% classification accuracy

within 30 epochs, compared with 73.6% using conventional preprocessing (Fig. 5e and Supplementary Fig. 19). Test results maintain accuracies above 90% across all illumination levels (Fig. 5f), demonstrating the advantage of high-sensitivity, ambipolar IKPs for illumination-robust in-sensor preprocessing.

2.5 Energy-Efficient Dynamic Perception

Beyond static feature extraction, the symmetric ambipolar photoresponse of the HGAFETs enables temporal processing, allowing the arrays to function as motion-aware IKPs. This capability is critical for edge-vision applications such as UAV surveillance, where high-bandwidth video streams (e.g., 178 MB s^{-1} for 1080p video at 30 fps) must be processed under strict constraints on latency, power, and computational resources. In typical dynamic scenes, moving objects occupy only a small fraction of each frame, rendering much of the spatiotemporal data redundant. Exploiting the symmetric NPC/PPC response, in-sensor frame differencing is realized by dynamically programming the V_{bg} of the HGAFET array (Fig. 6a). Consecutive frames are acquired alternately in NPC and PPC modes, with frame *N* recorded under NPC operation and frame *N*+1 under PPC operation. Due to the symmetry of the ambipolar photoresponsivity, static regions generate photocurrents of comparable magnitude but opposite polarity, which are eliminated through analog summation. In contrast, moving objects produce local intensity variations that generate strong differential signals, directly encoding motion information at the sensor output.

This functionality was experimentally demonstrated using a 32 × 32 HGAFET array (see the detailed in the Method and the optical image in Fig. 6b), with dynamic characterization performed on a representative 6 × 8 (48-pixel) subarray during a moving-vehicle scenario. For frame *N* containing only a static background, a positive V_{bg} (NPC mode) yields a uniform negative photocurrent map, with static pixels exhibiting currents of approximately −21.62 pA under low irradiance (Supplementary Fig. 20). In frame *N*+1, when the target enters the field of view, a negative V_{bg} (PPC mode) produces a positive photocurrent map. While static regions maintain a mean

photocurrent of 19.45 pA, target pixels exhibit a substantially enhanced response of 217.84 pA due to increased reflected illumination. Analog summation suppresses static background contributions while preserving motion-induced differential photocurrents, yielding high-contrast delineation of moving objects. A subsequent thresholding step converts the fused output into compact motion masks.

The resulting motion masks guide a lightweight YOLOv8 model, which processes only target-containing sub-images (Fig. 6c and Supplementary Fig. 21). Across 211 video frames, a data compression efficiency of 91.04% is achieved, corresponding to the retention of just 8.96% of the original pixel data. The average inference latency is reduced from 277.95 ms to 17.17 ms per frame (Fig. 6c, inset), supporting real-time UAV obstacle detection[46]. Compared with full-frame processing, this approach achieves a 16.18× reduction in latency and a 72.99× reduction in computational load, decreasing from 79.3 to 1.09 GFLOPs per frame (Fig. 6d and Supplementary Note 1). Benchmarking against representative motion perception architectures (Fig. 6e and Supplementary Note 2)[13,16,41,47,48] shows that the HGAFET-based in-sensor motion-aware IKPs achieve comprehensive advantage in balancing recognition accuracy, data reduction, and computational cost over other reported techniques. This advantage is enabled by the combination of high photosensitivity and a wide gate-tunable computational weight range, supporting flexible system-level optimization. The comprehensive advantage, along with demonstrated wafer-scale scalability and M3D integration compatible with Si ROICs, enable HGAFET arrays to act as efficient, low-latency, and energy-efficient IKPs for edge-vision applications under both day-night transitions and low-light conditions.

3 Conclusion

We develop a HGAFET consisting of an ambipolar field-effect transistor and a heterojunction gate, which enables the in-sensor IKPs extend to the shortwave infrared band. By leveraging the symmetric ambipolar transport in Schottky-barrier CNT FETs, the HGAFET converts a heterojunction-generated photovoltage into PPC or NPC

responses with the D^* exceeding 10^{13} Jones through adjusting the bottom-gate voltage. The direct physical mapping between photocurrent and V_{bg} enables precise in-sensor convolution for high-accuracy feature extraction under low illumination, while the symmetric ambipolar response supports analog-domain frame differencing for motion perception. Through integrating with Si ROICs based on M3D process, IKP arrays will provide a scalable hardware foundation for infrared edge-vision systems requiring high spatial resolution and energy efficiency, with potential applications in UAVs, micro-robots, and satellite-based imaging under ultralow-light and complex conditions.

4 Methods

Preparation of CNT Films

Arc-discharged single-walled CNT (SWCNT) powders were obtained from Carbon Solutions Inc. The SWCNTs were dispersed in toluene containing poly[9-(1-octylonoyl)-9H-carbazole-2,7-diyl] (PCz), followed by ultrasonication and gradient centrifugation (3,000 rpm, 30 min) to isolate high-purity semiconducting SWCNTs (99.99%). The resulting dispersion was diluted with fresh toluene to achieve the desired concentration. Pre-patterned substrates with local back-gate electrodes were immersed in the SWCNT dispersion at room temperature for 48 h to deposit randomly oriented CNT network films. The samples were then rinsed with pure toluene for 5 min, dried under high-purity nitrogen (99.999%), and annealed at 180 °C for 30 min to remove residual solvents and improve film adhesion.

Fabrication of p-i-n Heterostructures

The n-type ZnO layer was patterned on predefined channel regions using electron-beam lithography (EBL; Voyager, Raith) and room-temperature magnetron sputtering of a 40 nm ZnO film at 100 W. PbS QDs were synthesized via solution-phase ligand exchange following established protocols[40,49]. For the intrinsic layer, halide-passivated PbS QDs were spin-coated at 2,500 rpm for 20 s and annealed at 90 ℃ for 10 min, yielding ~150 nm thick films. The p-type layer was formed by sequential spin-coating of oleic-acid-capped PbS QDs (2,500 rpm, 20 s), followed by treatment with 0.02 vol% ethanedithiol

(EDT) in acetonitrile and rinsing with acetonitrile. This spin-coating/EDT-treatment cycle was repeated twice to produce ~40 nm p-doped QD films, completing the p-i-n heterostructure.

Fabrication of CNT FETs

CNT FETs were fabricated on Si/SiO_2 substrates with a 300 nm thermal oxide layer. Local back-gate electrodes (Ti/Au, 5/20 nm) were patterned by EBL, followed by electron-beam evaporation (EBE) and lift-off. A 15 nm HfO_2 gate dielectric was deposited by atomic layer deposition (ALD) at 150 ℃. Vias were defined by EBL and reactive ion etching (RIE) to electrically connect the local back-gate metal to the top probing pads. CNT network films were deposited as described above, and channel regions were defined by EBL and etched using RIE. Source and drain contacts (Ti/Au, 30/40 nm) were formed by EBE and lift-off. A 7-nm-thick HfO_2 film was deposited on the CNT channel by ALD, and the p-i-n heterostructure was integrated onto the HfO_2/CNT channel region as described previously.

Fabrication of Wafer-Scale HGAFET Arrays

Wafer-scale HGAFET arrays were fabricated on 4-inch Si/SiO_2 wafers using UV photolithography (EVG 610). Back-gate interconnects, a 15 nm HfO_2 dielectric, and via structures were defined prior to CNT deposition and patterning (Supplementary Fig. 14). Active regions were subsequently defined by photolithography and RIE. Source/drain electrodes and bottom interconnect metallization (Ti/Au, 30/40 nm) were deposited by EBE. An 80 nm SiO_2 layer was deposited by chemical vapor deposition to insulate the metal levels. Top-level interconnects (Ti/Au, 10/550 nm) were deposited, and complete arrays were wire-bonded to custom printed circuit boards (PCBs) for electrical readout.

Characterization and Measurements

CNT films were characterized by scanning electron microscopy (SEM; Zeiss Sigma 300) and atomic force microscopy (AFM; Bruker Dimension ICON). Electrical

transport measurements were performed on a probe station (MPI TS200-SE) using a semiconductor parameter analyzer (Keithley 4200). Optoelectronic measurements employed a supercontinuum laser source (NKT; 640–2100 nm) coupled to a monochromator (IHR320), with incident power density adjusted using calibrated attenuators (Thorlabs). Optical power was calibrated with certified Si (Thorlabs FDS100) and InGaAs photodiodes (Thorlabs FGA21), with calibrated intensities summarized in Supplementary Table 1-1 to 1-5. Beam profiles were recorded using an InGaAs camera (First Light Vision C-Red3). Noise spectral density was measured with a PDA FS380 system. Dynamic photoresponse and 3 dB bandwidth were recorded with an oscilloscope (Agilent DSO7054A).

Mechanism for Implementing MAC Operations

Under illumination, the p-i-n heterojunction generates photocarriers that induce a photovoltage V_{oc} (see Supplementary Note 3 for the explicit expression), which modulates the CNT channel operating in the subthreshold regime. The resulting drain–source current under illumination follows

$$logI_{ds,\ light} = \frac{\alpha V_{oc}}{SS} + logI_{ds,\ dark} \tag{2}$$

where α is the capacitive coupling coefficient describing the efficiency of photovoltage gating, SS is the subthreshold swing, and $I_{ds,\ dark}$ is the dark current. Substituting the expression for V_{oc} into Eq. (2) yields

$$I_{ds,\ light} = \left[\left(\frac{\eta q A P_{light}}{h\nu I_0} + 1\right)^{\frac{2.3\alpha k_B T}{qSS}}\right] I_{ds,\ dark}. \tag{3}$$

The photocurrent in the subthreshold regime is thus

$$I_{ph} = I_{ds,\ light} - I_{ds,\ dark} = \left[\left(aP_{light} + 1\right)^{b/SS} - 1\right] I_{ds,\ dark} \tag{4}$$

where $a = \eta qA/(h\nu I_0)$ and $b = 2.3\alpha k_B T/q$. The dark current itself follows the standard subthreshold relation

$$I_{ds,\ dark} = I_{0_FET} exp\left[\frac{q(V_{bg} - V_{th})}{nk_B T}\right] = Ce^{\beta V_{bg}}. \tag{5}$$

Here, I_{0_FET} is current factor of the CNT FET, V_{th} is the threshold voltage, n is the

slope factor, $C = I_{0_FET} exp[-qV_{th}/(nk_BT)]$, and $\beta = \frac{q}{nk_BT}$. Combining Eqs. (4) and (5), the photocurrent can be factorized into a light-dependent term and a gate-controlled weighting term: $I_{ph} = \left[\left(aP_{light} + 1\right)^{b/SS} - 1\right]\left(Ce^{\beta V_{bg}}\right) = I\left(P_{light}\right)W\left(V_{bg}\right)$. This multiplicative form directly enables analog matrix-vector multiplication in a HGAFET array: $I_{sum} = \sum_j I_j\left(P_{light}\right)W_j\left(V_{bg}\right)$. The physical separation between input encoding $I\left(P_{light}\right)$ and weight modulation $W\left(V_{bg}\right)$ establishes the foundation for in-sensor convolution and analog-domain computing using the HGAFET architecture.

Datasets and Experimental Evaluation

Low-light imaging experiments used grayscale images (8-bit, 0-255) converted from the Exclusively Dark Image Dataset[50]. Facial recognition experiments employed images from the Yale Face Database, comprising 165 images from 15 individuals. All images were downsampled to 64 × 64 pixels. Data augmentation expanded the original dataset to 1,320 images distributed across eight illumination levels. The augmented dataset contained 15 subjects, with 11 facial expressions per subject replicated under eight distinct illumination levels. Dataset diversity was quantified using grayscale histogram statistics (Supplementary Fig. 18).

Facial Recognition Task

A CNN framework was employed for facial recognition. The dataset was divided into training and test sets at a 7:3 ratio. Edge-enhanced outputs generated by in-sensor processing were fed into a CNN consisting of three convolutional layers (with 32, 64, and 128 channels) followed by two fully connected layers (with 256 and 15 neurons). ReLU activation functions were used throughout. Network weights were optimized using the Adam optimizer (learning rate = 0.0001) over 30 epochs by minimizing cross-entropy loss.

Motion-Aware Compression Algorithm

Motion-aware compression was experimentally implemented by dynamically

programming the back-gate voltage to encode opposite photoresponse polarities in consecutive frames. Position-dependent photocurrent mapping was performed on a representative 6 × 8 subarray extracted from the 32 × 32 HGAFET array, using synchronized Keithley 4200 source-measure units and an electronic control console with a spatial step size of 50 μm under 1300 nm illumination. For temporal frame differencing, the V_{bg} was alternated between consecutive frames: a positive V_{bg} was applied during Frame *N* to capture the static background (1.62 μW cm^{-2}), while a negative V_{bg} was applied during Frame *N*+1 to acquire the subsequent frame containing moving targets (3.61 mW cm^{-2}). Owing to the symmetric bidirectional photoresponse of the HGAFET, static regions generated photocurrents of nearly equal magnitude but opposite polarity in the two frames. Differential signal extraction was achieved by software-based summation of the positive photocurrent from Frame *N*+1 and the negative photocurrent from Frame *N*, effectively suppressing static background signals while preserving motion-induced differential responses. To evaluate algorithm-level performance beyond the hardware demonstration, a Python behavioral model was developed to emulate the motion-aware compression workflow using road-traffic video from the MATLAB® Computer Vision Toolbox™ sample library. Compression efficiency was assessed by feeding the resulting motion masks into a lightweight YOLOv8 object detection network and benchmarking detection performance against full-frame processing (Supplementary Note 1).

Supporting Information

Supporting Information is available from the XXXX or from the author.

Acknowledgements

We would like to thank Peking Nanofab for support in device fabrication and characterization. The authors acknowledge financial support from the National Science Foundation of China (Grant Nos. 62471028, 62225101, 22225202, and T2588301).

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

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

References

1. Almalioglu, Y., Turan, M., Trigoni, N. & Markham, A. Deep learning-based robust positioning for all-weather autonomous driving. *Nat. Mach. Intell.* **4,** 749–760 (2022).
2. Yeong, D. J., Velasco-Hernandez, G., Barry, J. & Walsh, J. Sensor and sensor fusion technology in autonomous vehicles: A review. *Sensors* **21,** 2140 (2021).
3. Ma, J., Ma, Y. & Li, C. Infrared and visible image fusion methods and applications: A survey. *Inform. Fusion* **45,** 153–178 (2019).
4. Chen, R. et al. Integrated bionic LiDAR for adaptive 4D machine vision. *Nat. Commun.* **17,** 24 (2025).
5. Oh, Y. et al. A 140 dB single-exposure dynamic-range CMOS image sensor with in-pixel DRAM capacitor. In *Proc. 2022 International Electron Devices Meeting 2022,* 37.7.1-37.7.4.
6. Lichtsteiner, P., Posch, C. & Delbruck, T. A 128 × 128 120 dB 15 μs latency asynchronous temporal contrast vision sensor. *IEEE J. Solid-State Circuits* **43,** 566–576 (2008).
7. Brandli, C., Berner, R., Yang, M., Liu, S.-C. & Delbruck, T. A 240 × 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor. *IEEE J. Solid-State Circuits* **49,** 2333–2341 (2014).
8. Posch, C., Matolin, D. & Wohlgenannt, R. A QVGA 143 dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression. In *Proc. 2010 IEEE International Solid-State Circuits,* 400–401.
9. Chai, Y. In-sensor computing for machine vision. *Nature* **579,** 32–33 (2020).
10. Zhou, F. & Chai, Y. Near-sensor and in-sensor computing. *Nat. Electron.* **3,** 664–671 (2020).

11. Wu, G. et al. Ferroelectric-defined reconfigurable homojunctions for in-memory sensing and computing. *Nat. Mater.* **22,** 1499–1506 (2023).

12. Ouyang, B. et al. Bioinspired in-sensor spectral adaptation for perceiving spectrally distinctive features. *Nat. Electron.* **7,** 705–713 (2024).

13. Huang, H. et al. In-sensor compressing via programmable optoelectronic sensors based on van der Waals heterostructures for intelligent machine vision. *Nat. Commun.* **16,** 3836 (2025).

14. Zhang, S. et al. Rational design of a vision fusion system with visible and near-infrared spectral integration for improved environmental perception. *Natl. Sci. Rev.* **12,** nwaf204 (2025).

15. Zhou, F. et al. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat. Nanotechnol.* **14,** 776–782 (2019).

16. Zhou, Y. et al. Computational event-driven vision sensors for in-sensor spiking neural networks. *Nat. Electron.* **6,** 870–878 (2023).

17. Lin, Q. et al. Event-driven retinomorphic photodiode with bio-plausible temporal dynamics. *Nat. Nanotechnol.* **20,** 1213–1220 (2025).

18. Yang, Y. et al. In-sensor dynamic computing for intelligent machine vision. *Nat. Electron.* **7,** 225–233 (2024).

19. He, Z. et al. In-pixel colour correction with organic self-adaptive transistors. *Nat. Photon.* **20,** 194–201 (2026).

20. Chen, J. et al. Optoelectronic graded neurons for bioinspired in-sensor motion perception. *Nat. Nanotechnol.* **18,** 882–888 (2023).

21. Gao, H. et al. Bio-inspired mid-infrared neuromorphic transistors for dynamic trajectory perception using $PdSe_2$/pentacene heterostructure. *Nat. Commun.* **16,** 5241 (2025).

22. Mennel, L. et al. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579,** 62–66 (2020).

23. Jang, H. et al. In-sensor optoelectronic computing using electrostatically doped silicon. *Nat. Electron.* **5,** 519–525 (2022).

24. Zhou, S. et al. Opto-electrical decoupled phototransistor for starlight detection. *Adv. Mater.* **37,** 2413247 (2025).

25. Zeng, L.-H. et al. Multilayered $PdSe_2$/perovskite Schottky junction for fast, self-powered, polarization-sensitive, broadband photodetectors, and image sensor

application. *Adv. Sci.* **6,** 1901134 (2019).

26. Han, Y. et al. Bipolar modulation in a self-powered ultra-wide photodetector based on Bi_2Se_3/AlInAsSb heterojunction for wavelength-sensitive imaging and encrypted optical communication. *Adv. Mater.* **37,** 2416935 (2025).
27. Xu, L. et al. Ultrasensitive dim-light neuromorphic vision sensing via momentum-conserved reconfigurable van der Waals heterostructure. *Nat. Commun.* **15,** 9011 (2024).
28. Fu, J. et al. Bionic visual-audio photodetectors with in-sensor perception and preprocessing. *Sci. Adv.* **10,** eadk8199 (2024).
29. Guo, T. et al. Reconfigurable phototransistors driven by gate-dependent carrier modulation in WSe_2/Ta_2NiSe_5 van der Waals heterojunctions. *ACS Nano* **19,** 1302–1315 (2025).
30. He, W., Wu, D., Kong, L., Yu, P. & Yang, G. Giant negative photoresponse in van der Waals Graphene/$AgBiP_2Se_6$/Graphene trilayer heterostructures. *Adv. Mater.* **36,** 2312541 (2024).
31. Ghosh, S. et al. Polarity-tunable photocurrent through band alignment engineering in a high-speed WSe_2/$SnSe_2$ diode with large negative responsivity. *ACS Nano* **16,** 4578–4587 (2022).
32. Du, J. et al. Gate-controlled polarity-reversible photodiodes with ambipolar 2D semiconductors. *Adv. Funct. Mater.* **31,** 2007559 (2021).
33. Jiang, J. et al. Wavelength-controlled photoconductance polarity switching via harnessing defects in doped $PdSe_2$ for artificial synaptic features. *Small* **20,** 2306068 (2024).
34. Ma, J. et al. Vertical 1T'-WTe_2/WS_2 Schottky-barrier phototransistor with polarity-switching behavior. *Adv. Electron. Mater.* **10,** 2300672 (2024).
35. You, J. et al. All-optic logical operations based on the visible-near infrared bipolar optical response. *Adv. Sci.* **11,** 2404336 (2024).
36. Liao, F. et al. Bioinspired in-sensor visual adaptation for accurate perception. *Nat. Electron.* **5,** 84–91 (2022).
37. Wu, J.-Y. et al. Broadband MoS_2 field-effect phototransistors: Ultrasensitive visible-light photoresponse and negative infrared photoresponse. *Adv. Mater.* **30,** 1705880 (2018).
38. Lin, Y., Miao, X., Zhang, Y., Hu, Y. & Pan, Y. Wavelength-dependent bipolar

photoresponse enabled all-in-one optoelectronic logic gates on epitaxial GaSe/GaN 2D-3D hybrid heterojunctions. *Adv. Optical Mater.* **13,** 2500416 (2025).

39. Chen, Y. et al. Ultranarrow-bandgap small-molecule acceptor enables sensitive SWIR detection and dynamic upconversion imaging. *Sci. Adv.* **10,** eadm9631 (2024).
40. Liu, J. et al. A near-infrared colloidal quantum dot imager with monolithically integrated readout circuitry. *Nat. Electron.* **5,** 443–451 (2022).
41. Wu, G. et al. Opto-electrical decoupling of phototransistors via light-induced ferroelectric depolarization for in-sensor computing. *ACS Nano* **19,** 20980–20990 (2025).
42. Zhao, Y. et al. High-performance all-inorganic perovskite tandem photodetectors with Bi_2TeO_6 Layer enabled by enhanced dual pyro-phototronic effect for triple-encryption imaging sensing system with programmable logic gate. *Adv. Mater.* **37,** e2414649 (2025).
43. Jiang, H. et al. Simultaneous achieving negative photoconductivity response and volatile resistive switching in Cs_2CoCl_4 single crystals towards artificial optoelectronic synapse. *Light Sci. Appl.* **13,** 316 (2024).
44. Liu, T. et al. Optically controllable 2D material/complex oxide heterointerface. *Adv. Sci.* **7,** 2002393 (2020).
45. Pi, L. et al. Broadband convolutional processing using band-alignment-tunable heterostructures. *Nat. Electron.* **5,** 248–254 (2022).
46. Hachaj, T. Potential obstacle detection using RGB to depth image encoder-decoder network: Application to unmanned aerial vehicles. *Sensors* **22,** 6703 (2022).
47. Huang, P.-Y. et al. Neuro-inspired optical sensor array for high-accuracy static image recognition and dynamic trace extraction. *Nat. Commun.* **14,** 6736 (2023).
48. Wu, Y. et al. CMOS-compatible retinomorphic Si photodetector for motion detection. *Sci. China Inf. Sci.* **66,** 162401 (2023).
49. Xia, Y. et al. Facet control for trap-state suppression in colloidal quantum dot solids. *Adv. Funct. Mater.* **30,** 2000594 (2020).
50. Loh, Y. P. & Chan, C. S. Getting to know low-light images with the Exclusively Dark dataset. *Comput. Vis. Image. Und.* **178,** 30–42 (2019).

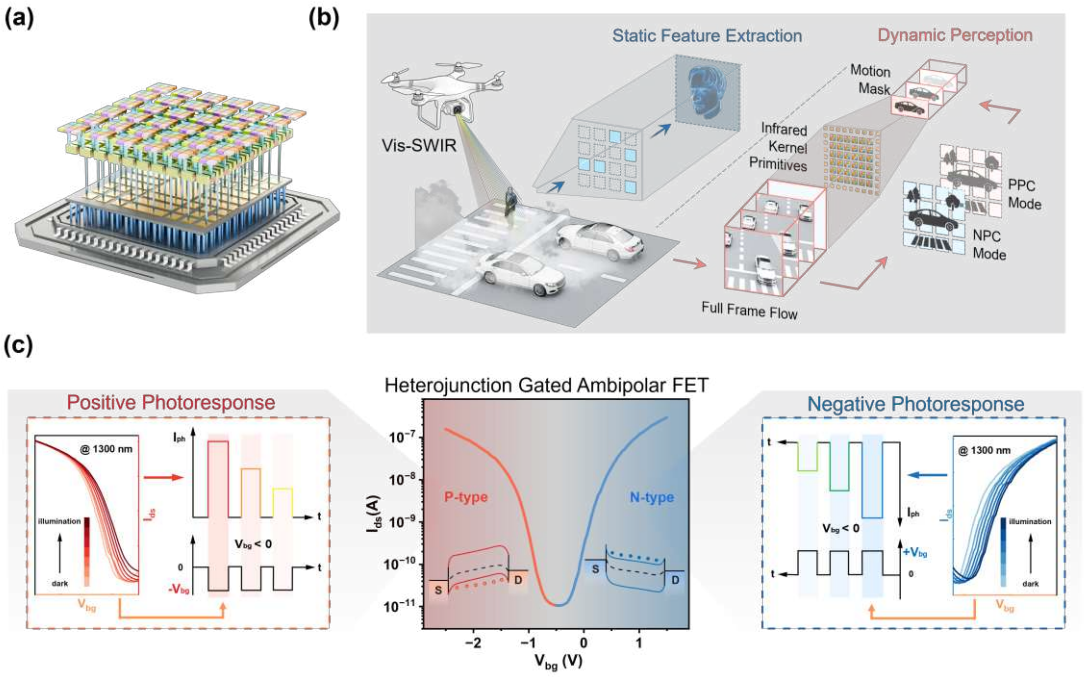


Fig. 1 | In-sensor IKPs based on HGAFETs. a, Schematic of an M3D image sensor integrating vertically stacked HGAFET arrays, memory, and computing units, providing a hardware platform for in-sensor IKPs. **b**, Broadband IKPs for static and dynamic perception. Static perception: Tunable HGAFETs modulate responsivity to extract facial edge features. Dynamic perception: Time-sequenced NPC and PPC biasing cancels static backgrounds via analog summation, allowing IKPs to generate a sparse motion mask. **c**, Working principle of a single HGAFET delivering symmetric, ambipolar, and gate-reconfigurable electrical outputs under broadband vis-SWIR illumination. Symmetric electron–hole injection is achieved in the SB FET. Insets illustrate the back-gate-tunable transition between PPC (left, $V_{bg} < 0$) and NPC (right, $V_{bg} > 0$), enabling analog in-sensor MAC operations and interframe-difference processing for both static and dynamic IKPs.

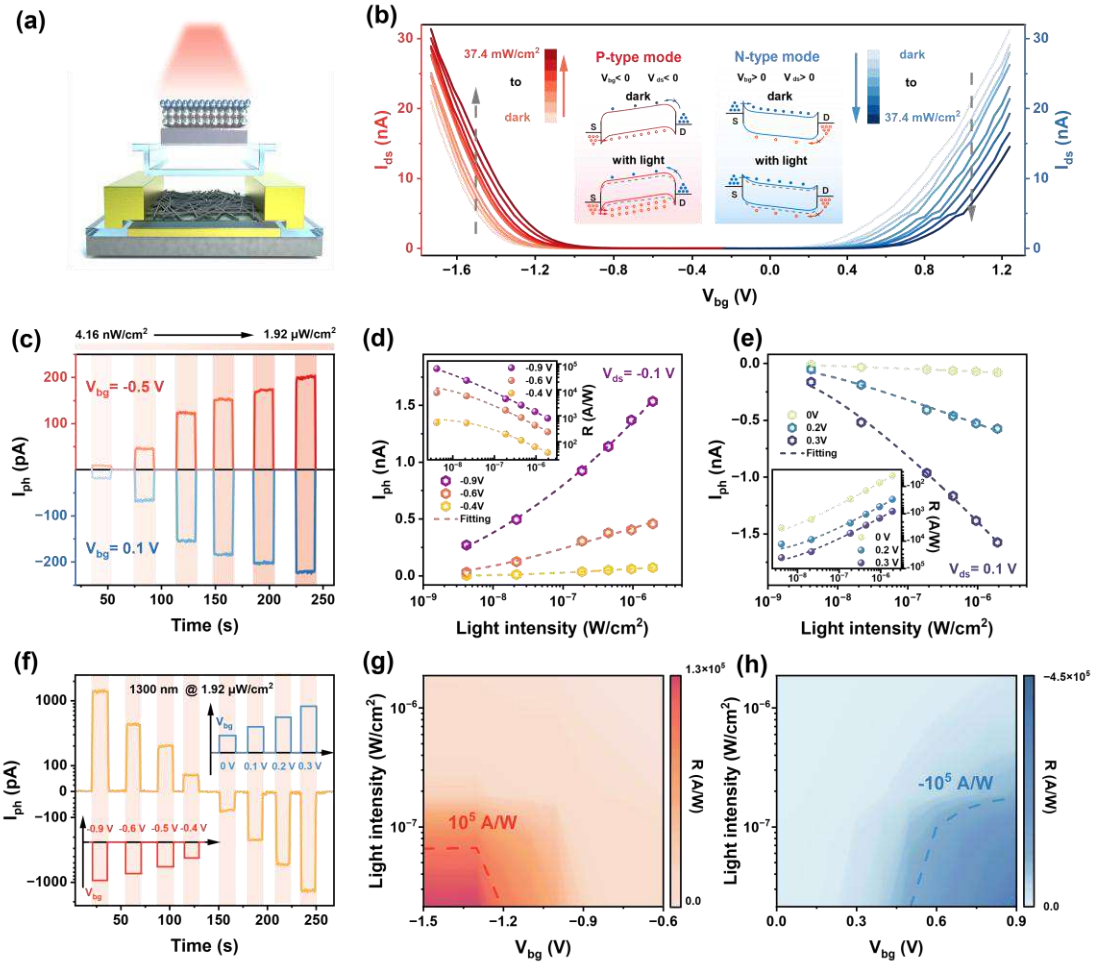


Fig. 2 | Gate-tunable symmetric and ambipolar photoresponse of HGAFETs under 1300 nm illumination. a, Device schematic showing the layer stack from top to bottom: PbS QD/ZnO heterostructure, HfO_2 top-gate dielectric, CNT channel, source/drain contacts, HfO_2 back-gate dielectric, and back-gate metal. **b**, Transfer characteristics in p-type and n-type regimes under increasing power density (0-37.4 mW cm^{-2}). Inset: schematic mechanism of ambipolar photoresponse. **c**, Time-resolved photoresponse at V_{bg} = -0.5 V (p-type) and 0.1 V (n-type) under ultralow irradiance (4.16 nW cm^{-2} to 1.92 μW cm^{-2}). **d**, **e**, Photocurrent as a function of incident light intensity at selected gate voltages: (**d**) p-region (V_{bg} = -0.4, -0.6, -0.9 V) and (**e**) n-region (V_{bg} = 0, 0.2, 0.3 V); dashed lines represent model fits. Insets show responsivity as a function of light intensity. **f**, Gate-controlled polarity switching of time-resolved photoresponse under 1.92 μW cm^{-2} illumination as V_{bg} is swept from -0.9 V to 0.3 V. **g**, **h**, Responsivity landscapes as a function of light intensity and V_{bg} in the (**g**) p-region and (**h**) n-region.

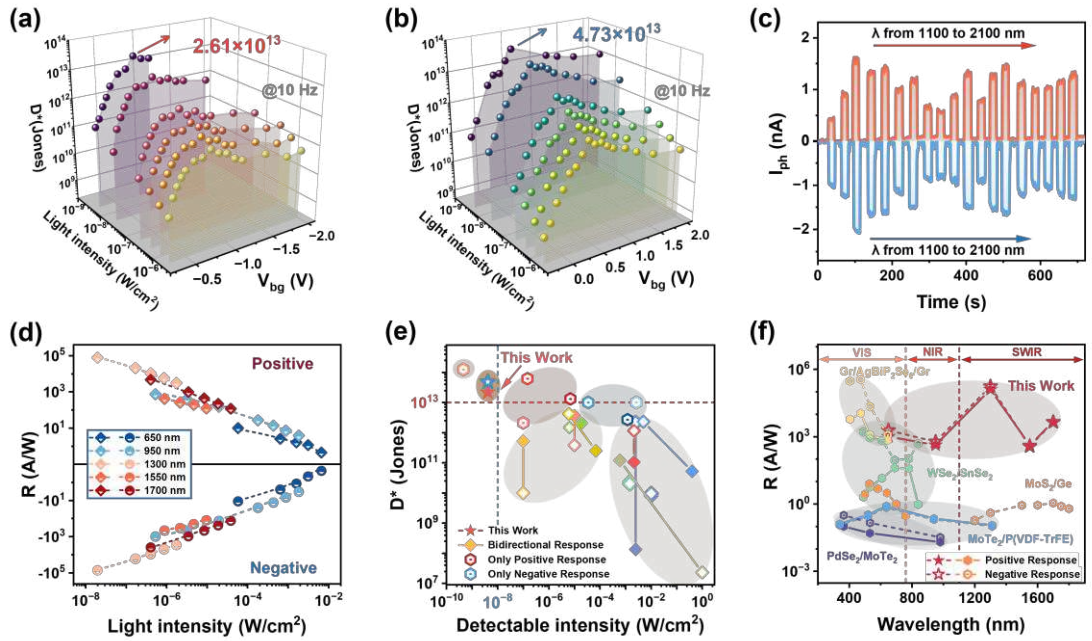


Fig. 3 | High detectivity and broadband ambipolar photoresponse. a, **b**, Specific detectivity (D^*) under (**a**) positive and (**b**) negative V_{bg} biases. **c**, Time-resolved NPC and PPC response across the vis-SWIR spectrum (1100-2100 nm, 40 nm steps) at V_{bg} = -0.8 V and 0.5 V; the 640-1080 nm range is shown in Supplementary Fig. 11. **d**, NPC and PPC responsivities at representative wavelengths (650, 950, 1300, 1550, and 1700 nm) under V_{bg} = 0.6 V and -1.2 V. **e**, Benchmark of minimum detectable light intensity vs. specific detectivity (D^*) for unipolar and ambipolar photodetectors (solid markers: positive response; hollow markers: negative response; see Supplementary Table 3). **f**, Responsivity comparison with other ambipolar photodetectors across wavelengths, highlighting the wide operational bandwidth and high responsivity of HGAFETs (see Supplementary Table 3).

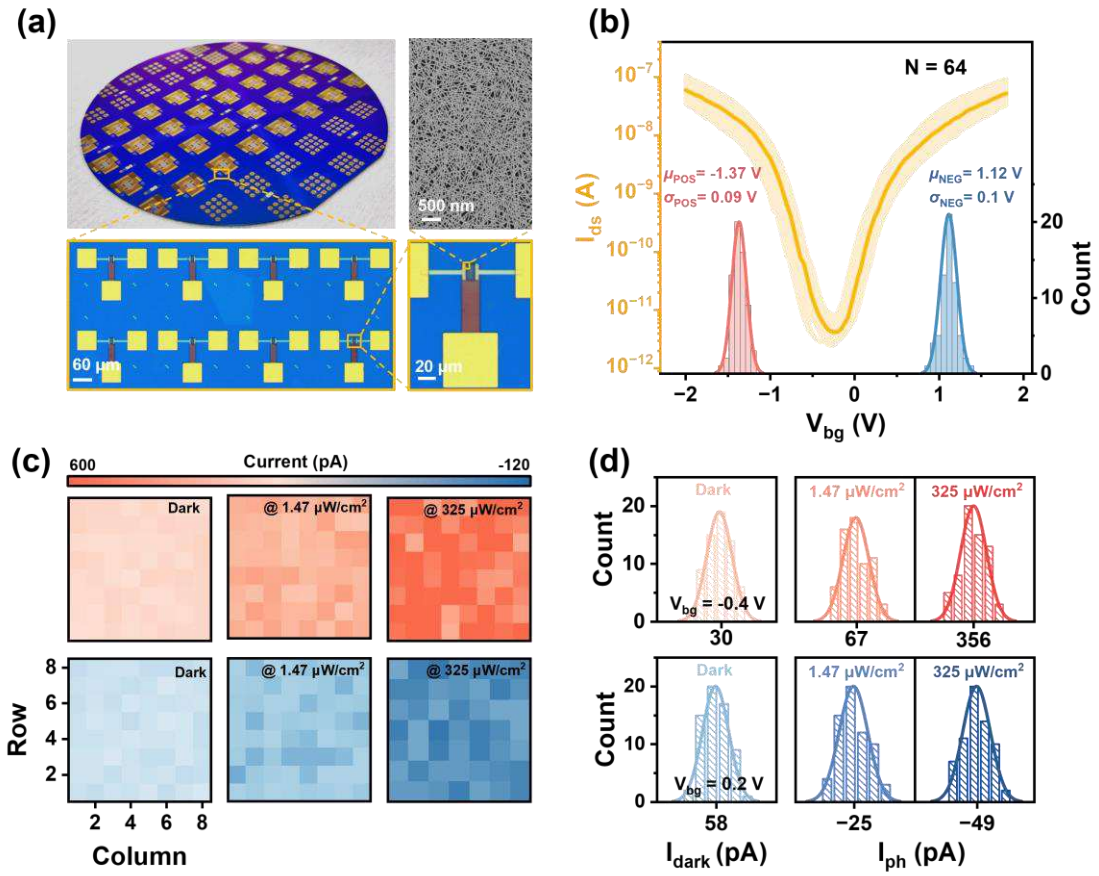


Fig. 4 | Wafer-scale CNT-based HGAFET arrays. a, Optical images of a 4-inch wafer (top-left) containing 28 complete dies, each with 1056 HGAFETs (bottom), and a high-resolution SEM image of the CNT channel. **b**, Ambipolar transfer curves for 64 randomly selected devices under dark conditions; inset shows statistical distributions of threshold voltages. **c**, Photocurrent maps under 1300 nm illumination at 0, 1.47, and 325 μW cm^{-2} with V_{bg} = -0.4 V and 0.2 V for PPC and NPC operation, respectively. **d**, Average and standard deviation of dark and photocurrent for the 64 devices in (**c**).

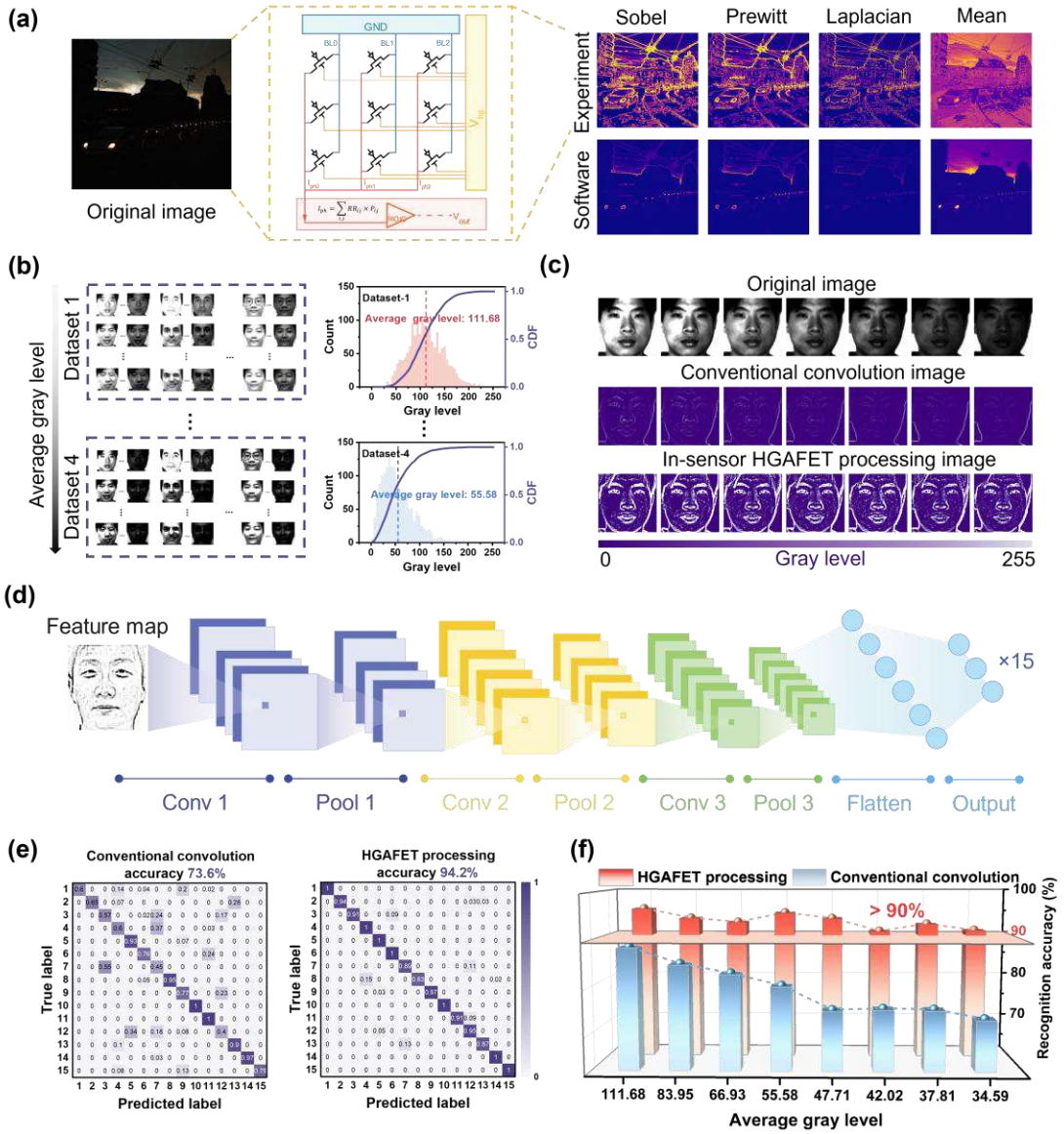


Fig. 5 | Low-light static image processing using IKPs. a, Schematic of in-sensor convolution implemented on a HGAFET array for edge and feature extraction, where gate-programmed weights map convolution kernels directly onto the device array. (Dataset source: Exclusively Dark Image Dataset) **b**, Augmented Yale Face Dataset with quantized average grayscale levels to simulate illumination variations. (Dataset source: Yale Face Database) **c**, Feature maps under varying illumination: conventional digital convolution (middle) versus IKP-based in-sensor preprocessing (bottom). **d**, Lightweight CNN framework integrating in-sensor IKP outputs for classification. **e**, Confusion matrices comparing conventional preprocessing and in-sensor feature extraction (15 classes). **f**, Recognition accuracy across eight illumination levels after 30 training epochs.

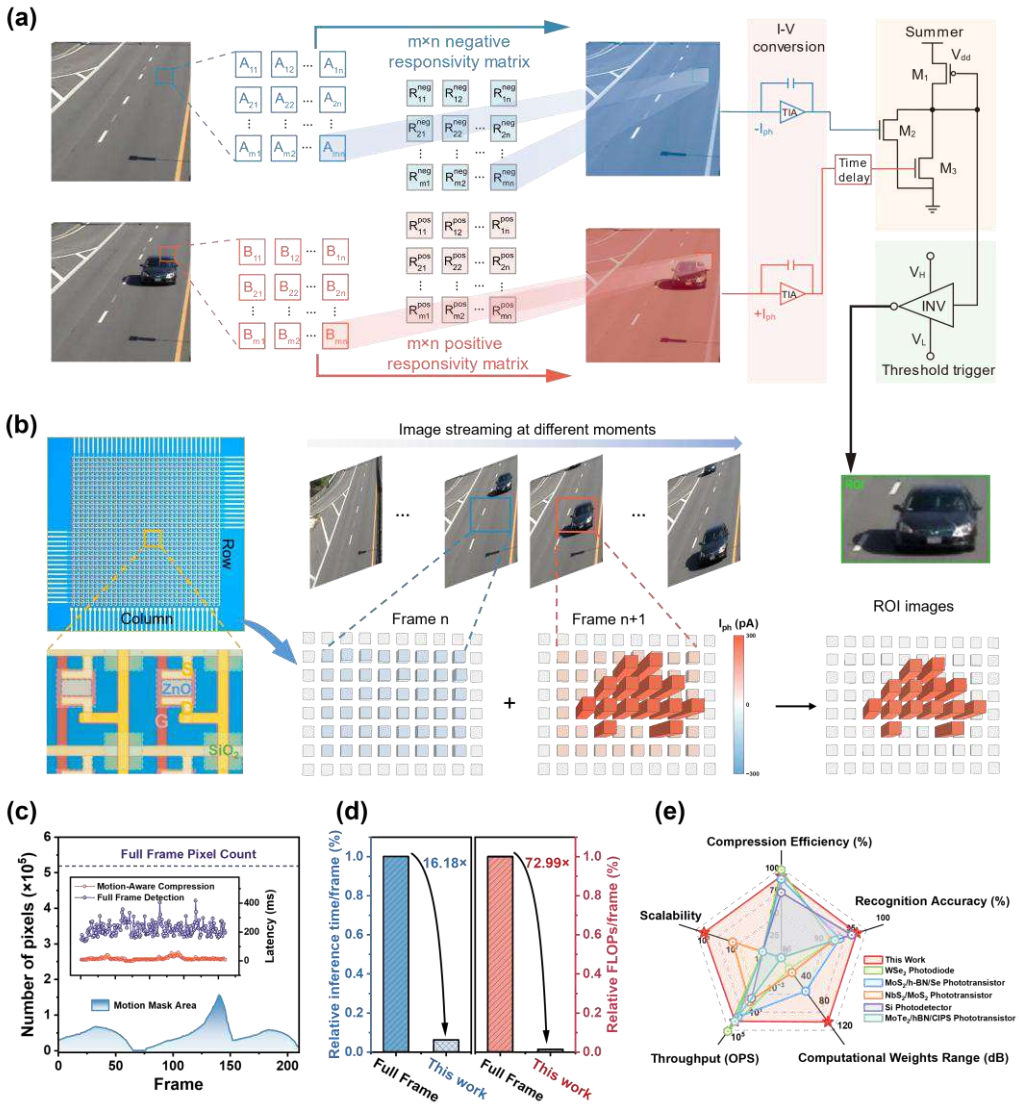


Fig. 6 | Energy-efficient dynamic perception using in-sensor IKPs. a, Schematic of in-sensor frame differencing enabled by ambipolar photoresponse in HGAFETs, producing pixel-level motion masks while suppressing static background. **b**, Experimental demonstration of motion mask generation via analog differential photocurrents. Left: optical image of a 32 × 32 HGAFET array; right: 6 × 8 (48-pixel) subarray showing motion-induced differential photocurrents. **c**, Statistical analysis of 211 video frames; motion masks retain only 8.96% of full-frame pixels on average. Inset: comparison of average inference latency, demonstrating acceleration enabled by in-sensor IKPs. (Video: MATLAB® Computer Vision Toolbox™ sample library) **d**, Motion-aware data compression using IKPs reduces inference latency by 16.18× and computational cost by 72.99× per frame compared with full-frame processing. **e**, Benchmarking of IKP-based in-sensor motion-aware compression against representative architectures across five key metrics (Supplementary Note 2).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- QXYangSI260522.docx