# A Near-array Convolution Computing Scheme Based on $WSe_2$ Photodiode

**Lingling Cai**, **Kejie Huang**, **Ruibing Song**, **and Haibin Shen**
College of Information Science & Electronic Engineering, Zhejiang University, Hangzhou, China

2021 Photonics & Electromagnetics Research Symposium (PIERS) | 978-1-7281-7247-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/PIERS53385.2021.9694838

**Abstract**— Due to the separation of perception and computing, the movement of data between perception devices and computing devices has become a bottleneck in intelligent systems. In that case, In-Sensor Computing (ISC) systems bring favorable opportunities for power-limited recognition devices. It integrates perception and computation, shortening the moving path of data, and therefore reducing energy consumption. However, the conventional schemes are still suffering from the issues of either low fill factor or low computing efficiency. Recently, a novel photodiode is developed using $WSe_2$ material, whose photoresponsivity can be controlled by the gate voltage. In this paper, an efficient ISC scheme is proposed for convolution operation to take full advantage of the above photodiode. In our scheme, the kernel weights can be stored in a continuously tunable photoresponsivity matrix. The proposed ISC scheme can realize flexible-sized kernel-readout with minimum one step for convolutional operations. A simulation is conducted on a $7 * 7$ pixel array, with specifically filter size of $3 * 3$, strides of 2, and padding of 0 (the realization of other array sizes and filter sizes will also be discussed). Meanwhile, the architecture is optimized to enable parallel convolution operation. The multiplication is achieved during the exposure period by preset the photoresponsivity matrix. The results show that the proposed ISC can achieve up to 20.6 TOPS/W energy efficiency with 9.71 pJ/pixel/frame power consumption and a maximum processing speed of 20 MHz, and the fill factor is 97%, outperforming the state-of-the-art.

## 1. INTRODUCTION

Developing imaging devices for the Internet of Everything (IoE), Intelligent Terminal, and Artificial Intelligence (AI), energy-efficient intelligent image sensors have become highly desirable in recent years. Some perceptual computing systems have been designed to execute learning algorithms with sensory data from the real-world environment. Among varieties of Neural Network (NN) algorithms, Convolution Neural Network (CNN) has shown significant improvement in image recognition and classification. Nevertheless, CNNs are computation-intensive and memory hungry, which challenges the convolutional integrated CMOS Image Sensor, making CNN tasks difficult to be deployed to energy-limited devices.

Many efforts in digital signal processing have been carried out to balance the performance and power of IoT devices, such as (GPUs), Field-Programmable Gate Arrays (FPGAs), and Neural Processing Units (NPUs) are used as hardware accelerators [1]. Till now, these digital solutions have not yet achieved satisfactory and practical improvement in efficiency. One of the reasons is that, as the most critical part of CNNs, convolution operations still consume considerable cost in the area, energy, and latency, not compatible with the well-optimized intelligent digital processor.

In recent years, Analog Signal Processing (ASP) [2] has aroused interest among researchers for its high energy efficiency. In Sensor Computing (ISC) [3] and Near Sensor Computing (NSC) [4] architectures are proposed, which focus on MAC (Multiply-Add Computation) operations [5] in analog domain to reduce the read-out consumptions and data density. In 2016, R. Li Kam Wa proposed a new architecture named RedEye [6], which transferred data pre-processing to the analog field, where multi-layer convolutional neural network operations can be performed. However, it costs extra area for analog computing circuits, resulting in a lower fill factor. In 2020, Zhe proposed an Processing Near Sensor Architecture (PNSA) [7], which can directly convert digital signal after analog calculation without Analog to Digital Conversion (ADC), but it only supports Binary Neural Network (BNN) which has low computing accuracy.

This paper proposed a near-array convolution computing scheme. This new scheme is based on a reconfigurable $WSe_2$ photodiode [8] array. The new type of photodiode's photoresponsivity can be linearly adjusted by gate voltage, and the multiplication operations can be implemented by encoding synaptic weights into the photoresponsivity matrix with no extra circuits. The simulations show that the ISC scheme can achieve up to 20.6 TOPS/W energy efficiency with 9.71 pJ/pixel/frame power consumption and a maximum processing speed of 20 MHz, and the fill factor is 97%, outperforming the state-of-the-art.

## 2. PROPOSED ARCHITECTURE

In this section, the proposed scheme is introduced in this order: (a) The pixel level circuit design with the $WSe_2$ photodiode to enable MAC operation; (b) the overall architecture and workflow of the scheme;

### 2.1. Pixel Circuit Design

Figure 1(a) shows the internal structure of a pixel unit. The basic unit is mainly composed of a PMOS transistor and four identical sets of photodetectors based on the new two-dimensional material $WSe_2$. The back gate voltage VG regulates the photosensitivity $R$, but the gate is only energized when the weight is preset. The relevant equations of the device are:

$$I_p = R \cdot E \cdot A = R \cdot P = R \cdot \frac{\frac{d\phi}{dS}}{k_\lambda} \cdot A \tag{1}$$

$$R = 0.3V_G \tag{2}$$

$$I_{\text{total}} = I_{\text{SAT}} \left( e^{\frac{qV_A}{K_B T}} - 1 \right) - I_p \text{ [9]} \tag{3}$$

where $I_p$ is photocurrent, $R$ is photoresponsivity, $E$ and $P$ denote irradiance and optical power respectively, $A$ is the detector area, $\frac{d\phi}{dS}$ is illuminance, $k_\lambda$ is conversion coefficient between luminous flux and optical power, $I_{\text{total}}$ is total current, $I_{\text{SAT}}$ is reverse saturation current, $q$ is electron charge, $V_A$ is bias voltage, $k_B$ is boltzmann constant, $T$ is temperature.

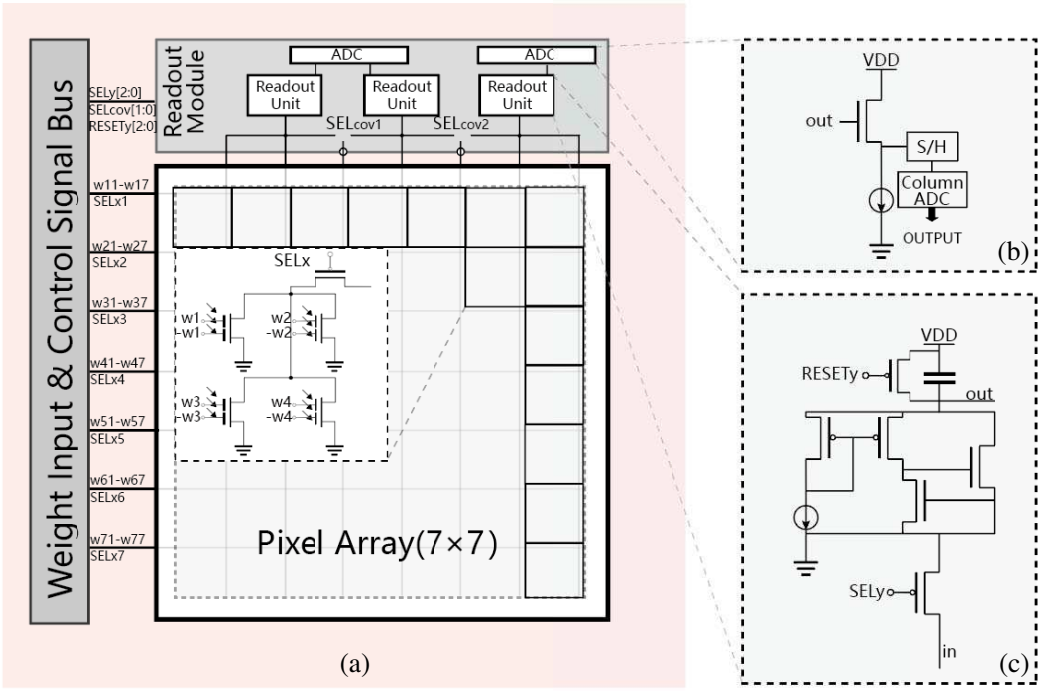



Figure 1: The structure of proposed ISC scheme.

The structure of the photodetector is the same as that of a floating gate transistor with a split-back gate, Fig. 2 shows the structure of the device. The gate is suspended during the exposure and convolution step, and the photodetector works in a short-circuit state. The PMOS transistor functions as a switch. Only when the PMOS transistor is gated, the pixel is activated.

### 2.2. The Overall Architecture and Workflow

Figure 1 is the overall architecture of the circuit, the size of the filter is $3 * 3$, the step is 2, the padding is 0, and the pixel array size is $7 * 7$. The whole scheme includes 49-pixel units, 196 photodetectors, 49 array strobe tubes, 2 switches, 3 read-out strobe tubes, 3 reset tubes, 100 fF capacitors, 3 voltage source, and 2 ADC modules. The pixels of each row share a row strobe signal SELx through the row bus. When SELx[$\text{Row}_m$] = on, it means that the pixel unit of the $m$th row

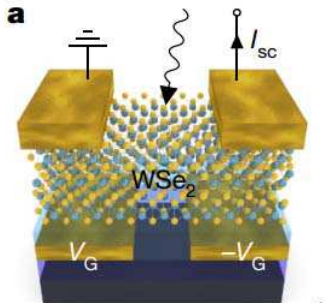



Figure 2: The structure of the device [8].

is selected, and all pixel units' switches $\text{Pixel}_{mi}$ $(i = 1, 2, 3, ..., 7)$ in the selected row are all set to the ON state. Meanwhile, a read-out unit $n$ is set outside the column of every $2n+1$ $(n = 1, 2, 3...)$ column of pixels, which is connected to the $2n-1, 2n, 2n+1$ column of pixels. In the $2m+1$ $(m = 1, 2, 3...)$, the connection between the column and the convolution module is provided with a 2-choose gating switch, because the column is connected to two read-out units. Every two read-out units share an ADC module, and they cannot perform readout operations at the same time.

Figure 1(c) is the transistor level readout unit, which is composed of the gate switch SELy, a rectifier module REGU, a large capacitor C, and the reset tube RESETy of the capacitor. The module REGU consists of a current source, two PMOS transistors and two NMOS transistors. It has little effect on small-scale pixel circuits but can be reduced Integral Nonlinearity (INL) caused by multiple parallel currents on larger-scale circuits. The large capacitor C and the reset tube RESET constructed by a PMOS tube are the core working components of the convolution unit circuit. During the reset phase, the gate of the PMOS receives a low-level strobe signal, the strobe tube is turned on, and the strobe tube RESET Pin 1 is connected to a 1.1 V voltage source, and pin 2 is connected to pin 2 of capacitor C. If the voltage of the bottom plate of capacitor C is less than 1.1 V, RESET will charge C. Within 10 ns, capacitor C will be fully charged. After that, in the process of convolution calculation, the gate signal of RESET is high, the PMOS tube is turned off. The capacitor C only discharges due to the photocurrent, and each discharge is always for a same duration, and the voltage change of the bottom plate caused by the discharge is read-out output of one convolution operation, and the data is processed in the digital domain after passing through the ADC module.

Figure 1(b) is the ADC module. In this research, only the general source-follower is used as the readout circuit.

The timing diagram of the controller is shown in Fig. 3, where 1 means that the strobe tube is turned on, and 0 means that the strobe tube is turned off.

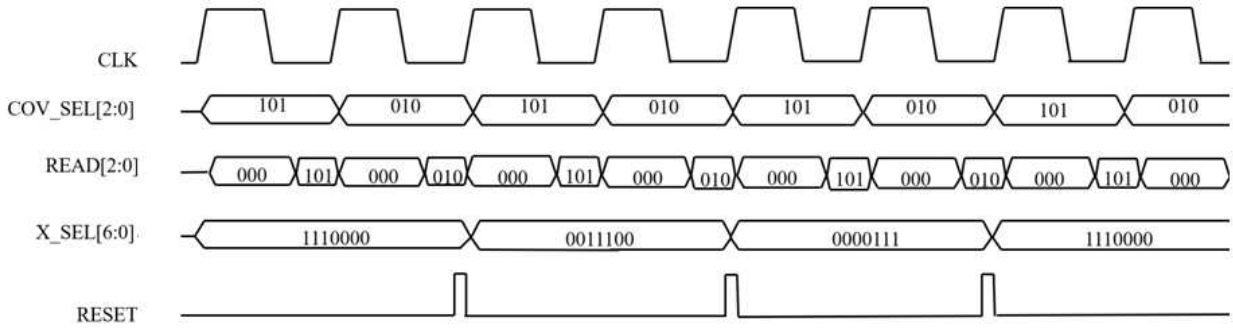



Figure 3: The timing diagram of the controller.

Since the capacitor takes into account the limitation of the highest voltage during read-out, the capacitor can only be discharged, not charged. Therefore, the positive weight and the negative weight are calculated separately during convolution. Fig. 3 only experiments with the positive weight value, and the negative convolution calculation can be repeated once more with the above

timing. For the circuit of this simulation, when the weight values of all convolution kernels are positive, this scheme only needs 6 clock cycles to complete all convolution calculations. In the 1st to 2nd clock cycles, the row strobe signal is 1110000, which means that all pixel units in the first row to the third row are strobed, and the photocurrent generated by it can flow into each column bus. In clock cycle 1, the strobe signal of the convolution unit is 101, which represents that the first and third convolution units are strobed, and the weight of the convolution kernel of the corresponding area is placed in each pixel unit. In clock cycle 2, the second convolution unit is gated, and the convolution kernel weight of the corresponding area is placed in each pixel unit; in the 3rd to 4th clock cycles, the row strobe signal is 0011100, which represents the 3rd. The 5 rows of pixel units are strobed (because the moving step is 2), the convolution operation sequence of the first row is repeated, and so on, until the convolution calculation of all pixels is completed, the convolution operation ends. It is worth noting that, in order to save energy, the readout circuit will be turned on at the end of each convolution operation, and the voltage value of the convolution result will be recorded. At the end of the convolution operation of a row, the reset tube is turned on to quickly reset all convolution units in the entire row. The schematic diagram of the convolution process is shown in Fig. 4.

Figure 4: The schematic diagram of the convolution process.

In this paper, the weight value $w_{ij}$ of the convolution kernel and the convolved input matrix $[x_{mn}]$ are respectively mapped to the back gate voltage $V_{G_{ij}}$ and the illuminance $\left(\frac{d\phi}{dS}\right)_{mn}$, and the expressions are as follows:

$$w_{ij} = C_1 V_{G_{ij}} \quad (i,j = 1,2,3) \tag{4}$$

$$x_{mn} = C_2 \frac{\frac{d\phi}{dS}_{mn}}{k_\lambda} \cdot A \quad (m,n = 1,2,3,4,5,6,7) \tag{5}$$

Among them, $C_1$ and $C_2$ are constant constants. In our simulation of the convolution, the illumination is constant, the input value $x_{mn}$ in this paper can be regarded as a two-value matrix, and the input value $x_{mn} = 1$ corresponding to the illuminated pixel, otherwise it is 0. As mentioned

above, the photosensitivity of the photodiode device of the photosensitive material of $WSe_2$ is controlled by the back gate voltage VG1 and VG2, and VG1 = −VG2. This article adopts a photodiode with a floating gate structure, which is preset at every resetstep.

After setting the weight value and input value, this article starts the convolution operation.

The controller of this circuit has a total of 4 sets of output signals, namely SEL_X[6:0], COV_SEL [2:0], READ, RESET, which control the row strobe signal, convolution strobe switch signal, and readout circuit selection respectively. The on/off switch signal and the reset tube switch strobe signal, the duration of the reset signal is 10 ns, the clock period is 6 μs, the duration of the readout circuit is 0.8 μs, and the duration of the convolution and row strobe signals is 5 μs.

During the cold start phase, the entire array is reset and initialized during startup, the reset tube is turned on, and other signals are in the off state, the 1.1 V power supply voltage source charges the capacitors of all convolution units until the bottom plate reaches 1.1 V. The weight is preset into the pixel unit to be convolved; in the exposure and convolution phases, the row strobe signal is partially turned on, the convolution signal is partially turned on, the reset signal is turned off, and the readout signal is turned off. The row strobe signal is used to select the specific three row signals, and the specific convolution strobe tube is turned on at the same time. In order to improve the computational efficiency, this article reads all the convolution areas without repeated input units at one time. For the $3 * 3$ convolution sum, the convolution of one line in this article actually only needs to perform two convolution operations, each time the adjacent convolution units does not perform convolution at the same time. In the readout phase, the convolution signal is turned off, the readout signal is turned on, and the weight value is refreshed at the same time. After two convolution and reading operations, the reset signal is turned on, the capacitor voltage is reset for all convolution units, and then the convolution operation of the next line is performed, and so on, until all convolution calculations are completed. When the weight value has a negative value, because it can only discharge the capacitor, so the convolution needs to be carried out twice, once the convolution of the positive weight, and the absolute value of the negative weight immediately after the convolution of a row. The convolution of the ADC is converted and subtracted to obtain the result of the convolution kernel convolution.

Assuming that a certain convolution unit is undergoing convolution operation at this time, the convolution time is the same each time, and the duration is $T$, then after one convolution is completed, the charge of the lower plate of the capacitor is:

$$Q_1 = CU_{\text{rst}} - I_{\text{total}}T = CU_{\text{rst}} - \sum_{i=1,\, j=1}^{I=9,\, J=4} \left(kw_{ij}x_{ij} + I_{\text{dark}}\right) T \tag{6}$$

Among them, $U_{\text{rst}}$ is the lower plate voltage of the capacitor before convolution, $C$ is capacitor, $w_{ij}$ is the weight value of the $j$-th photodetector of the $i$-th pixel unit, $I_{\text{dark}}$ represents the dark current, and $k = 1/C1$. Because the photodiode itself has dark current, and when the Vds is the same, it has nothing to do with the back-gate voltage VG and illumination. For this situation, this article will sample twice in a convolution process, and compare the positive part with the weight. The negative part is sampled:

$$Q_+ = CU_{\text{rst}} - \left[ CU_{\text{rst}} - \sum_{i=1,\, j=1,\, w_i>0}^{I=9,\, J=4} \left(kw_{ij}x_{ij} + I_{\text{dark}}\right) T \right] \tag{7}$$

$$Q_- = CU_{\text{rst}} - \left[ CU_{\text{rst}} - \sum_{i=1,\, j=1,\, w_i<0}^{I=9,\, J=4} \left(k|w_i|x_{ij} + I_{\text{dark}}\right) T \right] \tag{8}$$

By subtracting before and after, calculate the difference of the voltage, and get the result of a convolution:

$$\Delta Q = Q_+ - Q_- = \sum_{i=1,\, j=1}^{I=9,\, J=4} kw_i x_{ij} T \tag{9}$$

$$\Delta U = \frac{\Delta Q}{C} = \frac{1}{C} \sum_{i=1,\, j=1}^{I=9,\, J=4} kw_i x_{ij} T \tag{10}$$

If you want to get the convolution result of each positive weight and the convolution result of negative weight, this article defaults that the light condition does not affect the bias voltage of the photodiode, so this article only needs to collect in advance the volume of one convolution kernel under the non-light condition. The result of the product is the amount of charge change caused by the dark current (the current of the diode when there is no light):

$$Q_0 = CU_{\text{rst}} - \left( CU_{\text{rst}} - \sum_{0}^{36} (I_{\text{dark}})\, T \right) \tag{11}$$

After that, we can use its modified value to get the amount of change caused by the photocurrent each time:

$$\Delta Q_+ = Q_+ - Q_0 \tag{12}$$

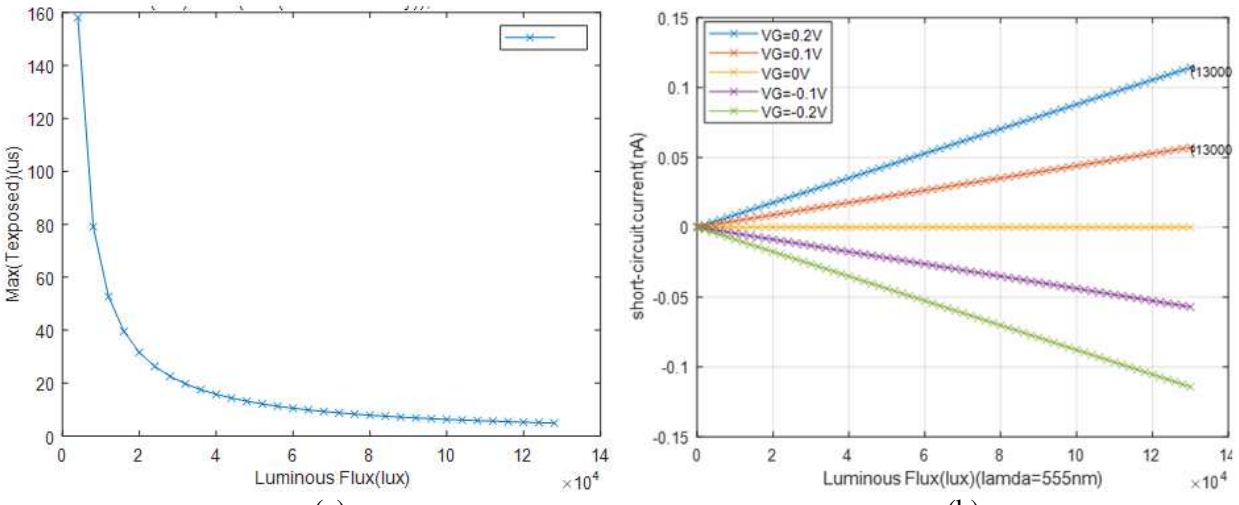
$$\Delta Q_- = Q_- - Q_0 \tag{13}$$

Since the voltage of the lower plate of the capacitor plays the function of recording the calculated value of the convolution, it also plays the function of driving the ADC output circuit. The driving voltage must be higher than a certain threshold, so the range of voltage change cannot be too large, otherwise the voltage is too low, cannot drive ADC output. We set the maximum voltage variation range of each exposure to $\Delta U_{\max}$. In order to withstand extreme conditions, we set the photocurrent of all photodiodes to the maximum sensitivity ($V_G = 0.2\,\text{V}$) to calculate the maximum exposure time:

$$\Delta T_{\text{expose}} = \frac{C\Delta U_{\max}}{I} = \frac{C\Delta U_{\max}}{\frac{R_{\max} * \phi * A}{K_{\lambda_{\max}}}} \tag{14}$$

Among them, $R_{\max} = \frac{0.06A}{W}$, $K_{\lambda_{\max}} = 683\,\text{lm/W}$.

If $\Delta U_{\max}$ is set to 0.2 V and $C$ is always 100 fF, the relationship between the longest exposure time and illuminance at this value is shown in Fig. 5. When the illuminance is 130000 lux, the maximum exposure time is less than 10 μs, but when the illuminance is less than 10000 lux, the maximum exposure time is above 100 μs. If we have higher requirements for the processing speed of the image sensor, we can properly preprocess the incident light to increase the illuminance of the photosensitive area. At the same time, monochromatic light is collected to improve working performance under low light conditions.




(a) (b)

Figure 5: (a) The relationship between illuminance and maximum exposure time ($|V_G| = 0.2\,\text{V}$, $\lambda = 555\,\text{nm}$); (b) The relationship between photocurrent and illuminance under typical back-gate regulation voltage.

For different illuminances, the exposure time varies greatly. In the case of higher illuminance, its working speed is faster. However, for some wavelength bands where the value of the visual function $K_\lambda$ is low, the longest exposure time will be shorter and the working rate will be relatively increased.

## 3. SIMULATION RESULTS

### 3.1. $WSe_2$ Photodiode Function Verification

The reverse current of the photodetector is extremely small, and the detected value is about $1.15 \times 10^{(-26)}$ A. We write verilog A code to perform behavior-level simulation of the photodetector in Virtuoso to ensure the device Works well in our sensor array. The simulation results are shown in Figure 5.

We fixed the wavelength and illuminance of the incident light, and performed DC simulation (Direct Current, DC) on the back gate voltage $V_G$. The simulation result is shown in Fig. 5. The visible light current and the back gate voltage $V_G$ are at $[-0.2\,\mathrm{V}, 0.2\,\mathrm{V}]$ linear relationship within the range.

### 3.2. Convolution Calculation Function Verification

In order to verify the function of the circuit, this experiment used Cadence Virtuoso 6.1.7 to build a simulation circuit for verification. This simulation focuses on the working state of the image sensor array. The simulation model adopted comes from the "5 nm Generic Process Design Kit" (GPDK45) provided by Cadence, with a 45 nm process and a logic level of 1.1 V/0 V. This simulation is based on a $7 * 7$ pixel unit array. We set the convolution kernel as a Sobel Kernel [10].

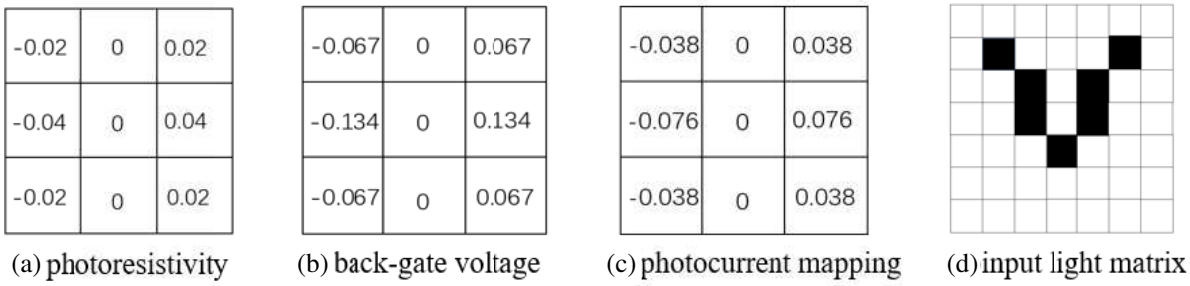



Figure 6: (a), (b), (c) The mapping values of photoresistivity, back-g at voltage and photocurrent mapping with sobel kernel; (d) light input matrix, black means active light input.

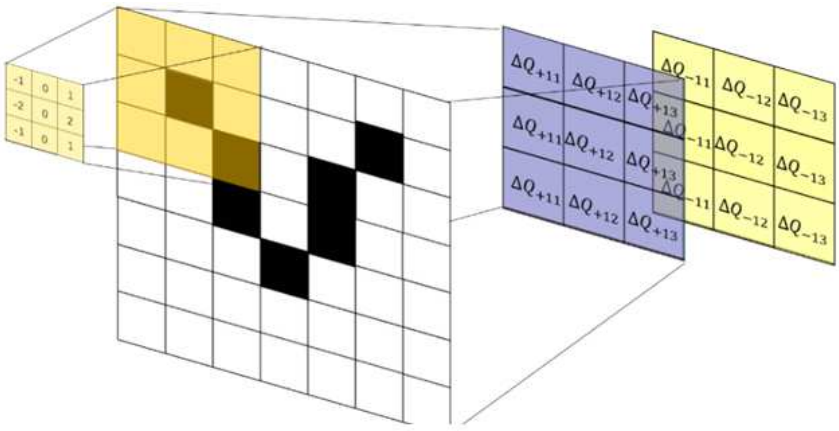



Figure 7: Diagram of the convolution process.

The weight values and input light matrix are shown in Fig. 6, and the process of the convolution are shown in Fig. 7. Among them, black represents the colored light area, and white represents the black non-light area.

A convolution sequence of a row is shown in Fig. 8. The clock cycle is 100 µs, the exposure time is 50 µs, the illumination is 130000 lux, and the wavelength is 555 nm. It can be seen from the above figure that the voltage change on the capacitor is basically linear with time. The voltage change value generated in the off state is about 0.003, $\Delta V_1 = V_1 - V_3$, $\Delta V_2 = V_2 - V_3$ can be obtained due to the photocurrent. Generated, through calculation, the effective average photocurrent is 0.038, which is in good agreement with the theoretical value, and the linearity is good.

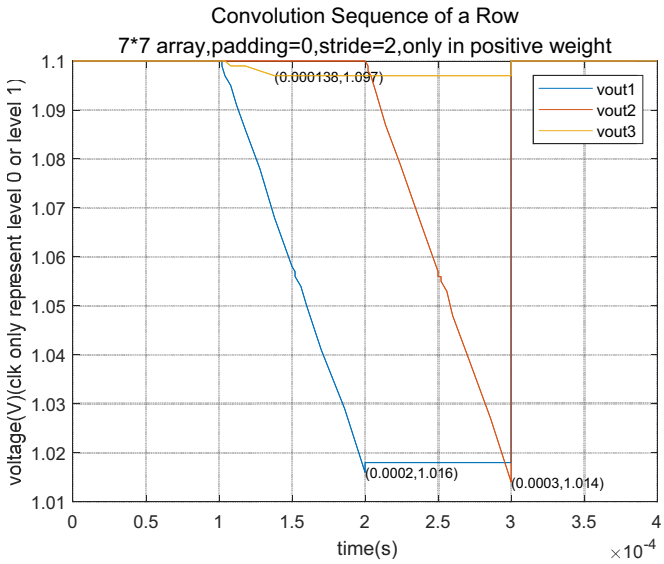



Figure 8: Covolution sequence of a row (with only posive weight considered).

### 3.3. Performance Analysis and Comparison of Related Work

PVT test's parameters is in Table 1, and the PVT test results is in Table 2; In Tables 1&2, the lowest point of the same curve has passed the test, and the error range is the same value; the same process angle is repeated three times, and the error range is the same value. There is an unavoidable jump at the moment when the convolution unit starts convolution (5 sets of process angles $*5$ sets of test currents $*5$ sets of different cycles $*3$ repeated tests, and the result is fitted to 0.00213). The results show that the relative error can range from 0.12% to 6.63% with different process corners and the mean value is 2.63%.

Table 1: The parameters of PVT test.

| *Process Corner* | *Temp.* | | | *Test Current* | *Test duration* |
|---|---|---|---|---|---|
| $Tt$ | −45 | 27 | 80 | 412 pA | 50 µs |
| $Ss$ | −45 | 27 | 80 | 412 pA | 50 µs |
| $fs$ | −45 | 27 | 80 | 412 pA | 50 µs |
| $sf$ | −45 | 27 | 80 | 412 pA | 50 µs |
| $ff$ | −45 | 27 | 80 | 412 pA | 50 µs |

Table 2: PVT test results.

| *process design kit* | *Process corner* | *Test Current* | *Test time* | $V_0$ (mV) | $V_1$ (mV) | $\Delta V$ | *Relative error* |
|---|---|---|---|---|---|---|---|
| gpdk045_v_3_0 | $tt$ | 412 pA | 50 µs | 1098.56 | 887.4757 | 211.0843 | 2.47% |
| gpdk045_v_3_0 | $ff$ | 412 pA | 50 µs | 1098.52 | 879.8381 | 218.6819 | 6.16% |
| gpdk045_v_3_0 | $fs$ | 412 pA | 50 µs | 1098.58 | 888.2716 | 210.3084 | 2.09% |
| gpdk045_v_3_0 | $sf$ | 412 pA | 50 µs | 1098.51 | 886.5292 | 211.9808 | 2.90% |
| gpdk045_v_3_0 | $ss$ | 412 pA | 50 µs | 1098.58 | 892.3393 | 206.2407 | 0.12% |
| gpdk045_v_3_0 | $mc$ | 412 pA | 50 µs | 1098.52 | 888.3383 | 210.1817 | 2.03% |
| Ideal value | | | | 1100 | 894 | 206 | |
| Mean | | | | 1098.545 | 887.132 | 211.413 | 2.63% |
| Standard deviation | | | | 0.029297 | 3.731673 | | |

The comparation with relative works is in Table 3. The readout power consumption of the image sensor array is reduced to 18.4% compared with the unconvolution calculation unit, and the power consumption of the capacitor charging is only 0.09% of the power consumption of the readout circuit, which can be compared with the traditional architecture. Improved a lot in power consumption. As for the delay, although the convolution calculation in the image sensor array requires a certain amount of time, its calculation efficiency can still be as high as 20.6 TOPS/W.

Table 3: Performance comparation.

| | *C. Shi et al. [11]* | *Bardallo et al. [12]* | *H. Zhu et al. [13]* | *Chen Z. et al. [3]* | *This work* |
|---|---|---|---|---|---|
| Architecture | APS, chip level design | APS pixel level design | DPS Pixel level design | APS Row-column design | PPS Pixel level design |
| convolution output | No | Yes | Yes | Yes | Yes |
| Process node | 180 nm | 350 nm | 65 nm | 180 nm | 45 nm |
| FoM | 6.41 nJ /frame/pixel | NA | 22.5 nJ /frame/pixel | 55.27 nJ /frame/pixel | 9.71 pJ /frame/pixel |
| Pixel complexity | 4 transistors | 49 transistors +3 capacitors | 46 transistors | 3 transistors | 1 transistor |
| Pixel area ($\mu m^2$) | $10 * 10$ | $40 * 40$ | $10.8 * 10.8$ | $40 * 40$ | $25 * 25$ |
| Fill Factor | 63% | 8.50% | 13.50% | 83.60% | 96.50% |

## 4. CONCLUSION

Based on a new type of photodetector model, this research proposes a sensor array architecture that can perform photoelectric conversion and convolution calculations in the analog domain. Through simulation verification and horizontal comparison, this work has achieved the expected functions on small-scale image sensors. It has excellent performance in pixel fill rate, power consumption (FoM), and computing efficiency (OP/S/W). The power consumption is greatly reduced. Finally, it is stated that all the content of this work is only implemented on Cadence, without considering non-ideal factors such as noise, parasitic capacitance, and parasitic resistance, and the reliability and stability of its work have not been tested by tape-out.

## REFERENCES


1. Lecun, Y., "Deep learning hardware: Past, present, and future," *2019 IEEE International Solid-State Circuits Conference-(ISSCC)*, 12–19, IEEE, 2019.
2. Yahampath, P., "Video coding for OFDM systems with imperfect CSI: A hybrid digital-analog approach," *Signal Processing: Image Communication*, Vol. 87, 2020.
3. Tyw, A., A. Jlm, A. Qxl, et al., "Reconfigurable optoelectronic memristor for in-sensor computing applications," *Nano Energy*, 2021.
4. Choi, S., K. Yong, T. V. Nguyen, et al., "Neural networks: Low-power self-rectifying memristive artificial neural network for near internet-of-things sensor computing," *Advanced Electronic Materials*, Vol. 7, 2021.
5. Chen, J., J. Li, Y. Li, et al., "Multiply accumulate operations in memristor crossbar arrays for analog computing," *Chinese Journal of Semiconductors*, Vol. 42, No. 1, 22, 2020.
6. Likamwa, R., Y. Hou, G. Yuan, M. Polansky, and Z. Lin, "RedEye: Analog ConvNet image sensor architecture for continuous mobile vision," *Computer Architecture News*, 2016.
7. Chen, Z., H. Zhu, E. Ren, et al., "Processing near sensor architecture in mixed-signal domain with CMOS image sensor of convolutional-Kernel-readout method," *IEEE Transactions on Circuits and Systems I: Regular Papers*, 1–12, 2019.
8. Mennel, L., J. Symonowicz, et al., "Ultrafast machine vision with 2D material neural network image sensors," *Nature*, 2020.

9. Koppens, F. H., R. T. Muelle, P. Avouris, et al., "Photodetectors based on graphene, other two-dimensionalmaterials and hybrid systems," *Nature Nanotechnology*, Vol. 9, No. 10, 780–793, 2014.
10. Wells, N. and C. W. See, "Polynomial edge reconstruction sensitivity, subpixel Sobel gradient kernel analysis," *Engineering Reports*, Vol. 2, 2020.
11. Shi, C., et al., "A 1000 fps vision chip based on a dynamically reconfigurable hybrid architecture comprising a PE array processor and selforganizing map neural network," *IEEE J. Solid-State Circuits*, Vol. 49, No. 9, 2067–2082, Sep. 2014.
12. Leñero-Bardallo, J. A., P. Häfliger, R. Carmona-Galán, and A. Rodríguez-Vázquez, "A bio-inspired vision sensor with dual operation and readout modes," *IEEE Sensors J.*, Vol. 16, No. 2, 317–330, Jan. 2016.
13. Zhu, H. and T. Shibata, "A real-time motion-feature-extraction image processor employing digital-pixel-sensor-based parallel architecture," *Proc. IEEE Int. Symp. Circuits Syst.*, 1612–1615, May 2012.