Received 24 December 2021; revised 8 February 2022; accepted 9 February 2022. Date of publication 11 February 2022;
date of current version 22 February 2022. The review of this article was arranged by Editor M.-D. Ker.

Digital Object Identifier 10.1109/JEDS.2022.3150922

Improved Crossbar Array Architecture for Compensating Interconnection Resistance: Ferroelectric HZO-Based Synapse Case

MINGFENG TANG [1], XUEPENG ZHAN [1,2], AND JIEZHI CHEN [1] (Senior Member, IEEE)

1 School of Information Science and Engineering, Shandong University, Qingdao 266237, China
2 State Key Laboratory of High-End Server & Storage Technology, Jinan 250100, China

CORRESPONDING AUTHORS: X. ZHAN AND J. CHEN (e-mail: zhanxuepeng@sdu.edu.cn; chen.jiezhi@sdu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 91964105, Grant 62104134, and Grant 62034006; in part by the National Key Research and Development Program of China under Grant 2016YFA0201802; in part by the Natural Science Foundation of Shandong Province under Grant ZR2020JQ28, Grant ZR2020KF016, and Grant ZR2019LZH009; and in part by the Young Scholars Program of Shandong University.

ABSTRACT In-memory computing is a promising solution to break through the conventional von Neumann bottleneck. Owing to the low-power consumption, Si fabrication compatibility and fast switching speed, the HfZrOx (HZO)-based ferroelectric devices attract attention as artificial synapses. By using crossbar architectures, artificial synapse array can greatly speed up the efficiency for neuromorphic applications. However, interconnect resistance effect will cause a serious decrease in calculation accuracy. This paper proposes a crossbar array architecture with the HZO-based ferroelectric synapses, which can restore the current distortion caused by the interconnect resistance. At the device level, the synaptic potentiation and depression behaviors are achieved by adjusting the pulse duration. For the circuit level, the interconnect resistance can be significantly compensated. In neuromorphic computing, a high accuracy rate of 96% is realized, which can be further improved with the expansion of the array size. Our results provide a step towards the development of large-scale ferroelectric HZO-based neuromorphic devices.

INDEX TERMS Ferroelectric synapse, residual polarization, interconnect resistance, neurological system.

I. INTRODUCTION

In-memory computing, showing fast calculation speed, is regarded as a promising candidate to break through the von Neumann bottleneck [1], [2]. It uses the computer memory to perform calculation and storage in the same location, without transferring data on the data bus [3]–[5]. The calculation efficiency is therefore significantly improved. Many novel devices are involved to build resistance crossbar arrays for the applications of in-memory computing, including the resistive random-access memory (RRAM) [6], phase change memory (PCM) [7]–[8], ferroelectric random-access memory (FRAM) and so on. Zhang *et al.* evaluated the respective advantages and disadvantages of RRAM. It is pointed out that the poor consistency, poor repeatability, and uneven conductance changes would bring problems in neurosynaptic device [9]. Wouters *et al.* proposed a PCM-based resistive switching memories, but its application is limited by the high energy consumption [10]. The ferroelectric random-access memory gets widespread attention because of its low power consumption [11] and fast switching [12], which is ascribed to the unique ferroelectric polarization conversion mechanism [13]–[14]. Oh *et al.* achieved multiple polarization residual states through different pulse voltages and durations [15]. But reset pulses are required to initialize the polarizations in each operation process, which could lead to unnecessary power loss.

In the crossbar array architectures, the larger size is beneficial for implementing more complex and accurate neural network algorithms [8]. However, the interconnect resistance effect and sneak path will lead to huge extra power [16]–[18]. When comes to the deep neural network (DNN) calculation [19]–[21], this effect can cause severe problems like decreased calculation accuracy and convergence speed. The main method to reduce this deviation is to substitute its

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

equivalent resistance through Kirchhoff's equation, but it takes a lot of time to calculate each memory. Liao *et al.* [16] and Bao *et al.* [22] replaced Kirchhoff's equations with a new algorithm to improve computational efficiency. But it still requires data transmission on the data bus, which reduces the advantages of the in-memory computing. Wu *et al.* [17] eliminated the data transformation by improving the unit structure of the in-memory computing array, which also confronts the problems of large extra power consumption. In addition, with the scaling in the device sizes of ferroelectric synapses, the enhanced reliability issue is becoming an important factor for device performances, which should be well-considered in designing circuit architectures [23]–[25]. Therefore, a solution that can maintain power consumption and ensure the calculation rate is highly required.

In this paper, a new resistance crossbar array architecture based on the HZO ferroelectric device is proposed. By adjusting the duration of the voltage pulse, the multi-level residual polarization states are obtained in ferroelectric synapse. In the crossbar array, the interconnect resistance can be compensated by using the improved circuits, which avoids solving Kirchhoff's equation. When comes to the DNN calculation, the proposed method can restore the accuracy, which is more suitable for larger crossbar arrays. Our finding provides the possibility for the development of neuromorphic computing based on ferroelectric artificial synapse.

II. FERROELECTRIC SYNAPSE

The ferroelectric synapse is composed of a 30nm TiN as the bottom electrode (BE), a 10nm HZO ferroelectric layer and a 40nm TiN as the top electrode (TE) and the electrode size is 100μm*100μm. The BE was grown on a silicon substrate with 200nm SiO_2 by sputtering (Vnano VJC-450). HZO is produced by atomic layer deposition (ALD, MNT S100-L3S1) at 280°C. The samples are annealed at 500 °C by rapid thermal-anneal (RTA, LABSYS RTP-1200) for 30 seconds to form ferroelectric phase. The switching speed of HZO-based devices largely depends on the grain sizes and distributions in the ferroelectric film, which can be modulated by using different pulse voltages and pulse durations [26]–[27]. As shown in Fig. 1(a), the PV curves are extracted experimentally under different pulse durations from 1 μs to 100 μs. A triangle pulse longer than 25 μs can switch the polarization state of the ferroelectric synapse completely, while the pulse shorter than 5 μs cannot. This indicates multiple residual polarization states can be achieved by using different durations.

Fig. 1(b) shows the extraction method of residual polarization state. Partial domain switching of ferroelectric devices is required for emulating the biological synaptic features. In each cycle, a switching pulse (500ns) is used to change the polarization states and a read pulse (500ns) is used for the extracted conductance. The residual polarization is decided by the imposed switching pulses in history (Fig. 1(c)). Sufficient margins can compensate for cycle-to-cycle and device-to-device variations, which are beneficial for neuromorphic applications. The gradual decreasing and increasing current are shown in Fig. 1(d), corresponding to the potentiation and depression behaviors of biological synapses.

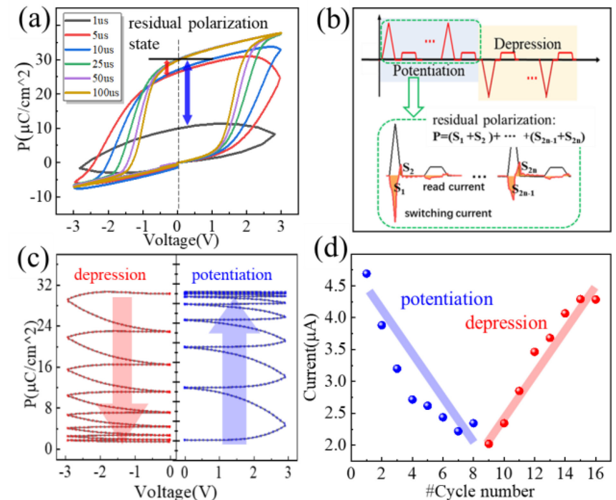

FIGURE 1. Device characteristics of single ferroelectric synapse. (a) Polarization voltage curves under different durations. (b) The scheme image residual polarization extraction method and whole test pattern. (c) Polarization switch process and (d) current change trend corresponding to the potentiation and depression process.

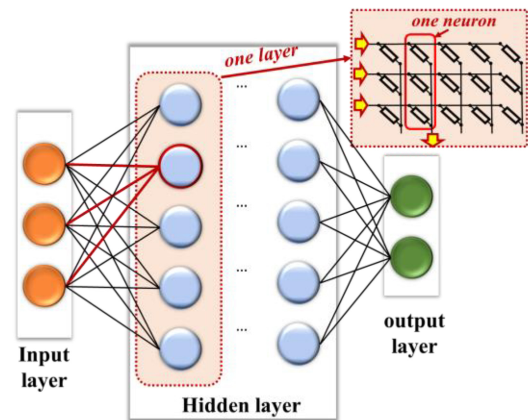

FIGURE 2. The diagram of DNN structure and the used memory array to represent a layer.

III. IMPROVED CIRCUITS METHOD

In ferroelectric synapse array, as shown in Fig. 2, a layer with m input signals and n neurons can be represented by a m × n array. The input signal is represented by voltage and transmitted to the next layer through current. It meets the requirements for the full connection of the previous layer and the next layer in the DNN calculation [28]–[29]. The output current obtained by the vector-matrix multiplication (VMM) is used as the input of the next layer, which greatly reduces the transmission of data in the data bus [18].

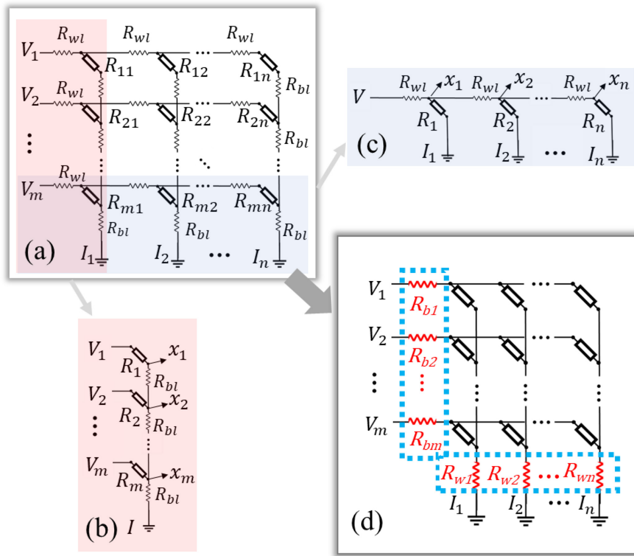

FIGURE 3. Resistance crossbar array circuit. (a) The equivalent circuit for solving Kirchhoff's equation. (b) The circuit of each column when ignoring the line resistance of the row circuit. (c) The circuit of each row when ignoring the line resistance of the column circuit. (d) The equivalent improved circuit.

In the resistance crossbar array, the ideal VMM calculation is:

$$I_j = \sum_{i=1}^{m} \frac{V_i}{R_i} \tag{1}$$

where I_j is the current output, V_i is the input voltage and R_i is the resistance of the memory. In the actual crossbar array, the interconnect resistance makes the calculated results suffering a large deviation. To reduce this error, a method to quickly calculate this effect is very necessary. Fig. 3(a) shows the equivalent circuit for solving Kirchhoff's equation. Fig. 3(b) and Fig. 3(c) consider the impact of each row or column separately, and Fig. 3(d) is a quick way to calculate this interconnect resistance effect by improved circuits.

When the effect on each column is ignored, as shown in Fig. 3(b), the effect of interconnect resistance on each row is the same. This means that only one line needs to be repaired and the result can be applied to every line in this array. The Kirchhoff equation at R_{bi} is:

$$\frac{x_{i-1} - x_i}{R_{bl}} + \frac{V_i - x_i}{R_{bi}} = \frac{x_i - x_{i+1}}{R_{bl}} \tag{2}$$

where V_i is the input voltage of each line, R_i is the resistance of the memory, and R_{bl} is the line resistance. When i is equal to 1, the first term in (2) does not exist. Thus, the formula becomes to:

$$\frac{V_1 - x_1}{R_1} = \frac{x_1 - x_2}{R_{bl}} \tag{3}$$

When the effect on each column resistance is ignored, two Kirchhoff equations can be expressed as:

$$\begin{bmatrix} \frac{V_1}{R_1} \\ \frac{V_2}{R_2} \\ \vdots \\ \frac{V_m}{R_m} \end{bmatrix} = \begin{bmatrix} \frac{1}{R_1} + \frac{1}{R_{bl}} & -\frac{1}{R_{bl}} & 0 & \cdots & 0 & 0 \\ -\frac{1}{R_{bl}} & \frac{1}{R_2} + \frac{2}{R_{bl}} & -\frac{1}{R_{bl}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{R_{bl}} & \frac{1}{R_m} + \frac{2}{R_{bl}} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \tag{4}$$

In the resistance crossbar array, the actual output current can be expressed as:

$$I_r = \sum_{i=1}^{m} \left(\frac{V_i - x_i}{R_{bi}} \right) \tag{5}$$

According to the difference between the ideal output I_j and the actual output I_r, the resistance of the additional circuit R_b can be obtained:

$$R_{bi} = \frac{mR_{bl}}{R_i + \sum_{i=1}^{m} (m+1-i)R_{bl}} \tag{6}$$

Using a similar analysis method, ignoring the effect on each row, as shown in Fig. 3(c), the interconnect resistance effect has the same effect on each column. The Kirchhoff equation at R_j is:

$$\frac{x_{j-1} - x_j}{R_{wl}} = \frac{x_j}{R_{wj}} + \frac{x_j - x_{j+1}}{R_{wl}} \tag{7}$$

where R_{wj} is the resistance of the memory, and R_{wl} is the line resistance. When j is equal to 1, x_{j-1} is the input V. When j is equal to n, the term $j+1$ does not exist, so Kirchhoff's equation can be written as:

$$\frac{x_{n-1} - x_n}{R_{wl}} = \frac{x_n}{R_{wj}} \tag{8}$$

The Kirchhoff equations when the resistance of the entire array is ignored can be expressed as:

$$\begin{bmatrix} \frac{V}{R_{wl}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{R_1} + \frac{2}{R_{wl}} & -\frac{1}{R_{wl}} & 0 & \cdots & 0 & 0 \\ -\frac{1}{R_{wl}} & \frac{1}{R_2} + \frac{2}{R_{wl}} & -\frac{1}{R_{wl}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{R_{wl}} & \frac{1}{R_n} + \frac{1}{R_{wl}} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{9}$$

In the same way, by making the difference between the ideal output current and the actual output current, the modified circuit can be obtained:

$$R_{wj} = \frac{\sum_{j=1}^{n} (j-n)R_{wl}}{R_j + \sum_{j=1}^{n} jR_{wl}} \tag{10}$$

In Fig. 3(d), two sets of circuits, R_b $(R_{b1}, R_{b2} \cdots R_{bn})$ and R_w $(R_{w1}, R_{w2} \cdots R_{wm})$, are added in the improved circuit. By using $R_{bi} + R_{ij} + R_{wj}$ to replace R_{ij}, the influence of interconnection resistance can be incorporated in the calculations. The proposed method can obtain similar results without solving the Kirchhoff's equations. From (6) and (10), the resistance of two addition circuits mainly depends on the line resistance and the position in the array.

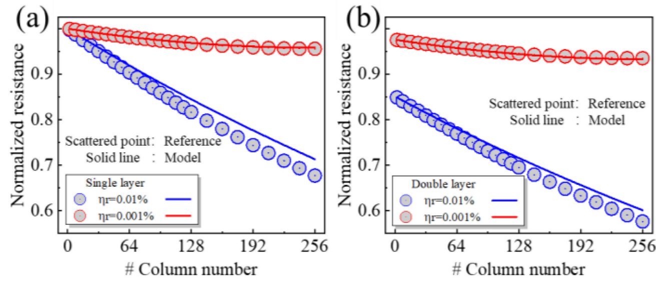

FIGURE 4. The comparison of equivalent resistance under different relative resistance obtained by using original circuit with solving Kirchhoff's equation and improved crossbar array circuit. (a) Single-layer array containing 784x256 devices; (b) Double-layer array containing 784x256 and 256x256 devices.

Fig. 4 displayed the comparison results of normalized resistance by using original circuit with solving Kirchhoff's equation (scattered point) and improved crossbar array circuit (solid line) in simulation. It can be seen that the output current has a large distortion in the array. The relative resistance (η_r) is defined as the ratio of interconnect resistance and device resistance. The interconnect resistance is depended by the wire widths and lengths. To simplify, the interconnect resistance is equal to the line resistance in each line and column ($R_{bl} = R_{wl}$). The device resistance is defined as the average of the highest and the lowest device resistance, which does not change in the training process. In the single-layer array (Fig. 4(a)), when the η_r is 0.001% and 0.01%, the normalized resistance has a distortion of 4.18% and 28.7%. This distortion will cause a significant decrease in the calculation accuracy of VMM operations. After using the improved circuit, the difference with the solution of Kirchhoff's equations is only 3.60%. This has reached acceptable accuracy. Incorporating this result into the calculation process can obtain the VMM calculation result more accurately. The compensation result of the double-layer array is shown in Fig. 4(b). Under different relative resistances, the improved circuit could compensate the influence of interconnection resistance, which is only 0.25% ($\eta_r = 0.001\%$) and 2.4% ($\eta_r = 0.01\%$) away from the result obtained by using original circuit with solving Kirchhoff's equation. For the multi-layer array, the proposed method can also improve the accuracy by incorporating the distortion into the calculation. Noted that this compensation method can be used in almost all the two-terminal resistance devices including FRAM, RRAM, PCM and so on.

IV. APPLICATION IN DNN CALCULATION

In DNN simulation, a ferroelectric synapse crossbar array with improved circuits is used to verify the compensation effect. The improved circuit is used to compensate errors in VMM calculations caused by interconnect resistance. To maintain the output balance of the resistance crossbar network, each layer of the network is composed of two circuits (Fig. 5(a)), leading to the weight of the

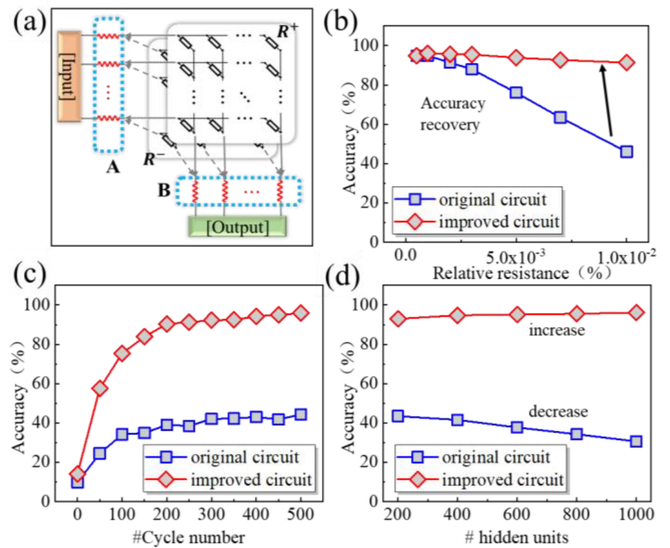

FIGURE 5. Application of ferroelectric synapse in DNN calculation. (a)Resistive crossbar array that can be used for DNN. (b)-(d) The accuracy recovery effect of the improved circuit.

output:

$$\text{output} = \frac{\text{input}}{\left(A + R^{+} + B\right) - \left(A + R^{-} + B\right)} \tag{11}$$

where *input* is the input voltage ($V_1, V_2 \cdots Vm$), *output* is the output current($I_1, I_2 \cdots I_n$), R^+ and R^- are the two differential circuits of the resistance crossbar array. The sum of all input signals across all modes of a particular class is ensured to zero, which speeds up the convergence process [30]. From (6) and (10), two additional circuits, A and B, can be obtained separately. Due to the array size is the same, this circuit can be used together to correct the output current. In the DNN calculations, different resistivities represent the weight of the layer, and are passed to the next layer by way of current.

The test is performed on the MNIST data through the DNN network [31]. This is a handwritten digit recognition data set consisting of 60,000 training sets and 10,000 test sets. Each data set is a 28×28 pixel image divided into ten categories. In Fig. 5(b), ignoring the impact of the interconnect resistance, recognition accuracy of 94.84% can be obtained with 1 hidden layer (150 units). With the interconnect resistance gradually increasing, the accuracy of the calculation decreases rapidly. Using the improved circuit, the accuracy of the calculation is maintained at a relatively high reference value. Even if the relative resistance reaches 0.01%, the calculation accuracy has reached 91.34%, which is much higher than the 45.89% accuracy by original circuits. It indicates that the proposed method can compensate interconnect resistance effect and restore the accuracy to a higher reference value.

The accuracy at different cycle numbers is shown in Fig. 5(c). With the improved circuit, the accuracy is improved by the same cycle number. Fig. 5(d) shows the accuracy at different unit numbers in the neural network. In the original

circuit, the accuracy slightly drops as the number increases due to the influence of the interconnect resistance. For the improved circuit, the accuracy increases with the increasing unit number. When it comes to 1000, an accuracy as high as 96.15% is obtained. Moreover, the impacts of device variation on the accuracy should be considered when the resistance distribution is larger than 5%.

V. CONCLUSION

In summary, a high-efficiency architecture is proposed in ferroelectric synapse crossbar arrays. By adjusting the duration of voltage pulse in the HZO based device, stable multi-level residual polarization states are achieved experimentally. At circuit level, the compensation effect of the improved circuit is theoretically verified. By the DNN algorithm, the calculation accuracy is restored to a higher reference value, which can be further improved in a larger array. Our findings provide possibilities for the development of neuromorphic applications based on ferroelectric synaptic crossbar arrays.

REFERENCES

[1] N. Verma *et al.*, "In-memory computing: Advances and prospects," *IEEE Solid-State Circuits Mag.*, vol. 11, no. 3, pp. 43–55, Aug. 2019.

[2] S. Yu, X. Sun, X. Peng, and S. Huang, "Compute-in-memory with emerging nonvolatile-memories: Challenges and prospects," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Boston, MA, USA, 2020, pp. 1–4.

[3] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," *Nano Lett.*, vol. 10, no. 4, pp. 1297–1301, 2010.

[4] D. Kuzum, S. Yu, and H. S. P. Wong, "Synaptic electronics: Materials, devices and applications," *Nanotechnology*, vol. 24, no. 38, 2013, Art. no. 382001.

[5] G. S. Rose *et al.*, "Design approaches for hybrid CMOS/molecular memory based on experimental device data," in *Proc. 16th ACM Great Lakes Symp. VLSI*, 2006, pp. 2–7.

[6] G. W. Burr *et al.*, "Experimental demonstration and tolerancing of a large-scale neural network (165 000 synapses) using phase-change memory as the synaptic weight element," *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3498–3507, Nov. 2015.

[7] J. Park, M. Kwak, K. Moon, J. Woo, D. Lee, and H. Hwang, "TiO_x-based RRAM synapse with 64-levels of conductance and symmetric conductance change by adopting a hybrid pulse scheme for neuromorphic computing," *IEEE Electron Device Lett.*, vol. 37, no. 12, pp. 1559–1562, Dec. 2016.

[8] G. W. Burr *et al.*, "Experimental demonstration and tolerancing of a large-scale neural network (165,000 synapses), using phase-change memory as the synaptic weight element," in *Proc. IEEE Int. Electron Devices Meeting*, 2014, pp. 1–4, doi: 10.1109/IEDM.2014.7047135.

[9] T. Zhang, K. Yang, X. Xu, Y. Cai, Y. Yang, and R. Huang, "Memristive devices and networks for brain-inspired computing," *Physica Status Solidi Rapid Res. Lett.*, vol. 13, no. 8, 2019, Art. no. 1900029.

[10] D. J. Wouters, R. Waser, and M. Wuttig, "Phase-change and redox-based resistive switching memories," *Proc. IEEE*, vol. 103, no. 8, pp. 1274–1288, Aug. 2015.

[11] S. Boyn *et al.*, "Learning through ferroelectric domain dynamics in solid-state synapses," *Nat. Commun.*, vol. 8, no. 1, pp. 1–7, 2017.

[12] Y. Nishitani, Y. Kaneko, M. Ueda, E. Fujji, and A. Tsujimura, "Dynamic observation of brain-like learning in a ferroelectric synapse device," *Jpn. J. Appl. Phys.*, vol. 52, no. 4S, p. 4CE06, 2013.

[13] J. Muller, T. S. Boscke, U. Schroder, R. Hoffmann, T. Mikolajick, and L. Frey, "Nanosecond polarization switching and long retention in a novel MFIS-FET based on ferroelectric HfO_2," *IEEE Electron Device Lett.*, vol. 33, no. 2, pp. 185–187, Feb. 2012.

[14] J. Müller *et al.*, "Ferroelectricity in HfO2 enables nonvolatile data storage in 28 nm HKMG," in *Proc. IEEE Symp. VLSI Technol. (VLSIT)*, 2012, pp. 25–26.

[15] S. Oh *et al.*, "$HfZrO_x$-based ferroelectric synapse device with 32 levels of conductance states for neuromorphic applications," *IEEE Electron Device Lett.*, vol. 38, no. 6, pp. 732–735, Jun. 2017.

[16] Y. Liao *et al.*, "Diagonal matrix regression layer: Training neural networks on resistive crossbars with interconnect resistance effect," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 40, no. 8, pp. 1662–1671, Aug. 2021.

[17] B. Wu *et al.*, "A low power reconfigurable memory architecture for complementary resistive switches," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 9, pp. 1806–1819, Sep. 2020.

[18] Y. J. Jeong, M. A. Zidan, and W. D. Lu, "Parasitic effect analysis in memristor-array-based neuromorphic systems," *IEEE Trans. Nanotechnol.*, vol. 17, no. 1, pp. 184–193, Jan. 2018.

[19] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognit. Lett.*, vol. 119, pp. 3–11, Mar. 2019.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 770–778.

[21] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*. Orlando, FL, USA: HRW Inc., 1987.

[22] L. Bao *et al.*, "Dual-gated MoS_2 neuristor for neuromorphic computing," *ACS Appl. Mater. Interfaces*, vol. 11, no. 44, pp. 41482–41489, 2019.

[23] K. Ni, W. Chakraborty, J. Smith, B. Grisafe, and S. Datta, "Fundamental understanding and control of device-to-device variation in deeply scaled ferroelectric FETs," in *Proc. Symp. VLSI Technol.*, 2019, pp. T40–T41.

[24] B. Sun *et al.*, "Synaptic devices based neuromorphic computing applications in artificial intelligence," *Mater. Today Phys.*, vol. 18, May 2021, Art. no. 100393.

[25] S. Yu, "Neuro-inspired computing with emerging nonvolatile memorys," *Proc. IEEE*, vol. 106, no. 2, pp. 260–285, Feb. 2018.

[26] C. Alessandri, P. Pandey, A. Abusleme, and A. Seabaugh, "Switching dynamics of ferroelectric Zr-doped HfO_2," *IEEE Electron Device Lett.*, vol. 39, no. 11, pp. 1780–1783, Nov. 2018.

[27] S. F. Lombardo *et al.*, "Local epitaxial-like templating effects and grain size distribution in atomic layer deposited $Hf_{0.5}Zr_{0.5}O_2$ thin film ferroelectric capacitors," *Appl. Phys. Lett.*, vol. 119, no. 9, 2021, Art. no. 92901.

[28] Q. Zheng *et al.*, "Artificial neural network based on doped HfO_2 ferroelectric capacitors with multilevel characteristics," *IEEE Electron Device Lett.*, vol. 40, no. 8, pp. 1309–1312, Aug. 2019, doi: 10.1109/LED.2019.2921737.

[29] Y. Sasago *et al.*, "Cross-point phase change memory with $4F^2$ cell size driven by low-contact-resistivity poly-Si diode," in *Proc. IEEE Symp. VLSI Technol.*, 2009, pp. 24–25.

[30] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-T. Müller, "Efficient BackProp," in *Neural Networks: Tricks of the Trade*. Heidelberg, Germany: Springer, 2012, pp. 9–48.

[31] A. F. Agarap, "An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification," 2017, *arXiv:1712.03541*.