Energy-Efficient Time-Domain Computation for Edge Devices: Challenges and Prospects

Other titles in Foundations and Trends® in Integrated Circuits and Systems

Recent Advances in Testing Techniques for AI Hardware Accelerators
Arjun Chaudhuri, Ching-Yuan Chen and Krishnendu Chakrabarty
ISBN: 978-1-63828-240-2

Of Brains and Computers
Jan M. Rabaey
ISBN: 978-1-63828-120-7

Emerging Trends of Biomedical Circuits and Systems
Mohamad Sawan, Jie Yang, Mahdi Tarkhan, Jinbo Chen, Minqing Wang, Chuanqing Wang, Fen Xia and Yun-Hsuan Chen
ISBN: 978-1-68083-906-7

Revisiting the Frontiers of Analog and Mixed-Signal Integrated Circuits Architectures and Techniques towards the future Internet of Everything (IoE) Applications
Rui P. Martins, Pui-In Mak, Sai-Weng Sin, Man-Kay Law, Yan Zhu, Yan Lu, Jun Yin, Chi-Hang Chan, Yong Chen, Ka-Fai Un, Mo Huang, Minglei Zhang, Yang Jiang and Wei-Han Yu
ISBN: 978-1-68083-892-3

Welcome to the World of Single-Slope Column-Level Analog-to-Digital Converters for CMOS Image Sensors
Albert Theuwissen and Guy Meynants
ISBN: 978-1-68083-812-1

Energy-Efficient Time-Domain Computation for Edge Devices: Challenges and Prospects

Hamza Al Maharmeh
Wayne State University
hamza.m@wayne.edu

Mohammed Ismail
Wayne State University
ismail@wayne.edu

Mohammad Alhawari
Wayne State University
alhawari@wayne.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Integrated Circuits and Systems

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

H. Al Maharmeh *et al.*. *Energy-Efficient Time-Domain Computation for Edge Devices: Challenges and Prospects*. Foundations and Trends® in Integrated Circuits and Systems, vol. 3, no. 1, pp. 1–50, 2024.

ISBN: 978-1-63828-357-7

Foundations and Trends® in Integrated Circuits and Systems
Volume 3, Issue 1, 2024

Editorial Board

Editorial Scope

Foundations and Trends® in Integrated Circuits and Systems survey and tutorial articles in the following topics:

- Analog, digital and mixed-signal circuits and systems
- RF and mm-wave integrated circuits and systems
- Wireless and wireline communication circuits and systems
- Data converters and frequency generation
- Power electronics and power management circuits
- Biomedical circuits and systems
- Sensor and imager circuits and cyber physical systems
- Security and resilient circuits and systems
- Circuits and systems in emerging non-CMOS technologies
- Circuit theory, modeling, analysis and design methods

Information for Librarians

Foundations and Trends® in Integrated Circuits and Systems, 2024, Volume 3, 4 issues. ISSN paper version 2693-9347. ISSN online version 2693-9355. Also available as a combined paper and online subscription.

Contents

Energy-Efficient Time-Domain Computation for Edge Devices: Challenges and Prospects

Hamza Al Maharmeh[1], Mohammed Ismail[2], and Mohammad Alhawari[3]

[1] *Wayne Center for Integrated Circuits and Systems (WINCAS), Wayne State University, USA; hamza.m@wayne.edu*
[2] *Wayne Center for Integrated Circuits and Systems (WINCAS), Wayne State University, USA; ismail@wayne.edu*
[3] *Wayne Center for Integrated Circuits and Systems (WINCAS), Wayne State University, USA; alhawari@wayne.edu*

ABSTRACT

The increasing demand for high performance and energy efficiency in Artificial Neural Networks (ANNs) and Deep Learning (DL) accelerators has driven a wide range of application-specific integrated circuits (ASICs). In recent years, this field has started to deviate from the conventional digital implementation of machine learning-based (ML) accelerators; instead, researchers have started to investigate implementation in the analog domain. This is due to two main reasons: (a) better performance, and (b) lower power consumption. Analog processing has become more efficient than its digital counterparts, especially for Deep Neural Networks (DNNs), partly because emerging analog memory technologies have enabled local storage and processing known as compute-in-memory (CIM), thereby reducing the

Hamza Al Maharmeh, Mohammed Ismail and Mohammad Alhawari (2024), "Energy-Efficient Time-Domain Computation for Edge Devices: Challenges and Prospects", Foundations and Trends® in Integrated Circuits and Systems: Vol. 3, No. 1, pp 1–50.
DOI: 10.1561/3500000013.

amount of data movement between the memory and the processor. However, there are a lot of challenges in the analog domain approach, such as the lack of a capable commercially available non-volatile analog memory, and the analog domain is susceptible to variation and noise. Additionally, analog cores involve digital-to-analog converters (DACs) and analog-to-digital converters (ADCs), which consume up to 64% of total power consumption. An emerging trend has been to employ time-domain (TD) circuits to implement the multiply-accumulate (MAC) operation. TD cores require time-to-digital converters (TDCs) and digital-to-time converters (DTCs).

However, DTC and TDC can be more energy and area efficient than DAC and ADC. TD accelerators leverage both digital and analog features, thereby enabling energy-efficient computing and scaling with complementary metal–oxide–semiconductor (CMOS) technology. The performance of TD accelerators can be substantially improved if custom-designed analog delay cells, DTC, and TDC are used. This work reviews state-of-the-art TD accelerators and discusses system considerations and hardware implementations. Additionally, the work analyzes the energy and area efficiency of the TD architectures, including spatially unrolled (SU) and recursive (REC) architectures, for varying input resolutions and network sizes to provide insight for designers into how to choose the appropriate TD approach for a particular application. Furthermore, it discusses our implemented scalable SU-TD accelerator synthesized in 65 nm CMOS technology with an efficient DTC circuit that utilizes a laddered inverter (LI) circuit that consumes 3× less power than the inverter-based DTC and achieves 116 TOPS/W. Finally, we discuss the limitations of time-domain computation and future work.

1

Introduction to Efficient Computing

1.1 Introduction

Deep Neural Networks (DNNs) have become the cornerstone for modern artificial intelligence (AI) applications due to the unprecedented achieved accuracy in image classification [47], [76], object recognition/detection [35], [77], [82], speech recognition [21], [22], [24], [40], [71], [74], [86], game playing [17], [65], [83], [84], healthcare [6], [30], [99], [102], [104], and robotics [51], [73], [81], [103]. As DNNs require significant computational resources, energy, and memory bandwidth to process huge amounts of data with small latency and high accuracy [88], they are typically implemented on the cloud using GPUs. Moving DNNs, however, out of the cloud into the edge devices provides key benefits, including improving privacy in some applications, such as healthcare, and reducing latency, which is critical in modern applications like autonomous driving.

Over the last 50 years, Moore's law and Dennard scaling have helped build faster, smaller, and energy-efficient transistors, but this trend has slowed down during the last decade due to the physical limits of the transistors [25], [96]. To overcome this limitation, various levels of research have built specialized computing hardware that can deliver high performance with high energy efficiency. Digital accelerators can

be custom-made specifically for DNNs and thus can provide higher throughput, shorter latency, lower energy, and higher area efficiency [9], [18], [19], [23], [26], [29], [43], [55], [66], [67], [87], [88], [92], [94]. Although digital accelerators provide better performance compared to GPUs, digital systems (including both GPUs and digital accelerators) are fundamentally limited in handling big data efficiently due to the separation of logic and memory (referred to as von Neumann bottleneck). Consequently, the system bandwidth is limited by the speed of accessing the data in the memory. Moreover, memory access requires at least 10x more energy/delay compared to the multiply-accumulate (MAC) operation [41], [88]. Hence, data movement in GPUs and digital accelerators dominates energy consumption and bandwidth.

To overcome the fundamental challenges in digital systems, analog and mixed-signal hardware accelerators have been explored to build artificial neural networks (ANNs) that can outperform the digital-based ones by several orders of magnitudes in energy efficiency, computation, and training time [11], [12], [16], [27], [28], [32], [45], [49], [54], [58], [75], [89], [91], [98], [100]. Analog computations promise simplicity and energy efficiency with real-time parallel processing and learning. Analog processing has become more efficient than the digital counterparts, especially for DNNs, partly because emerging analog memory technologies have enabled local storage and processing as shown in Figure 1.1(c), thereby reducing the amount of data movement between the memory and the processor. Besides, the MAC operations can be performed more efficiently in the analog domain. An ultimate example of analog computation is the human brain, which can perform more than 1016 operations/second while consuming 20 Watts [61]. Although analog computation is efficient in terms of energy and area [91], [98], it has limited accuracy and technology scaling [7]. Additionally, the need for Analog-to-Digital converters (ADCs) and Digital-to-Analog converters (DACs) limits the efficiency and scalability of the analog cores.

An emerging trend is to utilize time-domain (TD) to perform MAC operations by representing the data as pulses with modulation, as depicted in Figure 1.1(d) [3], [7], [20], [31], [56], [64], [79]. The goal is to perform the MAC operations in TD by producing proportional delays to inputs and weights. These delays are accumulated and then converted

back to digital. TD cores require time-to-digital converters (TDCs) and digital-to-time converters (DTCs). However, DTC and TDC can be more energy and area efficient than DAC and ADC, respectively [63]. Time-based accelerators can achieve superior performance while being energy efficient [2], [7], [20], [31], [56], [64], [79]. It has been shown that TD-ANN can achieve superior performance with excellent energy and hardware efficiency [2], [7], [20], [31], [56], [64], [79]. The digital approach has the best use of technology scaling, but it is not as efficient as the analog approach [7], [91], [98]. Time-based computation can take advantage of both approaches, analog and digital, as it is energy efficient and can be scaled with CMOS technology.

1.2 Background and Prior Work

The concept of neural networks was inspired by the biological neural system and was first conceived in 1943 [60]. Fully connected DNNs (FC-DNN) consist of multiple layers, an input layer that matches the width of the input data, an output layer that depends on the specific inference task and hidden layers. Figure 1.1(a) shows a feedforward FC-DNN architecture that is based on MAC or vector-matrix multiplication (VMM), where a vector of n neuron excitations, x_i, is multiplied by a vector of weights, w_{ij}, generating a new vector of neuron excitations for the next layer, y_j, and then followed by a nonlinear function, f.

$$y_i = f\left(\sum_i^n w_{ij} x_i\right). \tag{1.1}$$

In the digital domain, Figure 1.1(b) shows how a GPU implements the MAC operations, by using a large number of Arithmetic Logic Units (ALU) with the help of memory (DRAM) that stores the weights. Figure 1.1(c) shows an example of analog domain implementation where memory technologies have enabled local storage and processing. Figure 1.1(d) represents the basic implementation of TD neuron. As depicted in the figure, a chain of variable delay elements are cascaded where each element has a delay value that depends on the dot product between the corresponding input and weight. An input pulse is applied at

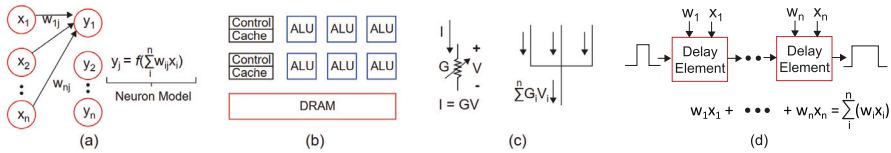

Figure 1.1: (a) DNN with basic mathematical operations, (b) GPU architecture, (c) analog computation, (d) time-domain computation.

the first element where the output pulse width will eventually represent the MAC result for the neuron.

1.3 Introduction to Analog Computing

To overcome the fundamental challenges in digital systems, analog and mixed-signal hardware accelerators have been explored to build artificial neural networks (ANNs) that can outperform the digital-based ones by several orders of magnitudes in energy efficiency, computation, and training time [11], [12], [16], [27], [28]. Analog computations promise simplicity and energy efficiency with real-time parallel processing and learning. Analog processing has become more efficient than its digital counterparts, especially for DNNs, partly because emerging analog memory technologies have enabled local storage and processing, thereby reducing the amount of data movement between the memory and the processor. Analog computing provides the ultimate in-memory computing (IMC) as it can be implemented in crossbar architecture. Besides, the MAC operations can be performed more efficiently in the analog domain. An ultimate example of analog computation is the human brain, which can perform more than 1016 operations/second while consuming 20 Watts [61]. The 2018 IBM Summit, one of the world's fastest supercomputers, may have a computing capacity comparable to that of the human brain [48], but it consumes 13 MegaWatts, with an area of two basketball courts. Hence, to explore the capabilities of future computing, the human brain can potentially offer design tricks to implement non-von Neumann architectures toward highly efficient and massively parallel computing platforms [62]. The advantages of analog computation are (a) its superior energy efficiency as it mitigates data

movement and memory access for the neural network weights, and (b) its extremely high throughput as the current passes through the PEs in every column of the crossbar to get the MAC output.

1.3.1 Hardware Implementation of Analog Computation

Analog hardware accelerators utilize crossbar-based architectures and emerging non-volatile memories (NVM), such as Resistive RAM (RRAM) [5], [10], [15], [42], [46], [52], [53], [72], [97], Phase-Change RAM (PCRAM) [8], [14], [33], [80], [95], and Magnetic RAM (MRAM) [13], are commonly used to build DNN systems. The crossbar architecture has rows and columns, where the NVM memory resides at the intersection between each row and column. This enables local storage and processing in a highly parallel and energy-efficient manner [11], [12], [16], [27], [28].

A key component in DNNs is the memory to store the value of the weights [44]. Analog memory technologies can be divided into two categories: charge-storage and non-charge-storage memories. Charge-storage memories depend on storing electric charges for an extended period of time. Floating-gate or embedded flash memory (eFM) is a NVM, charge-type memory that is used in DNNs [27], [39], [70]. eFM has a tunnel gate oxide at the channel interface, but due to the stringent requirement of long retention time, this tunnel oxide is already at its minimum thickness and is no longer scalable [13]. Furthermore, eFM requires high voltage pulses for programming and erasing, thereby potentially leading to high power consumption and long training times. Non-charge-storage memories are typically two-terminal, NVM devices, including RRAM, PCRAM, and MRAM.

1.3.2 Limitations of Analog Accelerators

In addition to the storage elements challenges in IMC architecture mentioned previously, analog accelerators are sensitive to noise, not like their digital counterpart which deals with two levels; 0 s and 1 s. The major sources of noise are thermal noise from electronic devices and quantization noise that comes from the crossbar interface circuits which are the data converters that convert the data from digital to analog and then from analog to digital [37], [68]. The issue becomes more

challenging for higher precision as it is well known that the thermal noise is proportional to $\sqrt{kT/C}$, and thus to achieve an extra bit, the capacitance needs to be increased fourfold and as a result, the energy will be increased 4 times [68]. So it is very challenging to design a high-precision analog core while being more efficient than digital cores. It has been shown that analog accelerators can be more energy-efficient than digital for low-bit precision, i.e., below 6–7 bits, otherwise digital will outperform the analog core [37], [68], [69], [78], [93].

In IMC architecture, the cost of accumulation operation is directly related to the conversion from analog to digital using an analog-to-digital converter (ADC). One ADC can be used for each column at the crossbar architecture, or multiple columns can share a high-speed ADC by using the means of reusing and time-multiplexing. The energy per MAC operation can be expressed as follows [68]

$$E_{\text{MAC}} = E_{\text{ADC}}/N + E_{\text{CAP}} + E_{\text{Logic}} \tag{1.2}$$

Where E_{ADC} is the ADC's conversion energy, and N is the number of rows. E_{CAP} and E_{Logic} are the energy consumption due to the unit capacitances (C_u) and logic gates in each processing element. It has been shown that for higher bit precision (greater than 7 bits), the ADC energy will dominate [34], [37], [68]. Other works reported that the ADC consumes 64% of total energy in [59], and 50% of total core power in [36], which urges the need for energy and area-efficient ADC designs.

1.4 Digital vs. Analog vs. Time-Domain Computing

The adaptable nature of the digital implementation allows for scalability using CMOS technology. However, due to data representation as a multi-bit digital vector, as the number of bits increases, so does the quantity of MAC units and operations. This leads to heightened dynamic switching capacitance, resulting in increased power consumption and additional area overhead [7], [9], [18], [19], [23], [26], [29], [43], [55], [66], [67], [87], [88], [92], [94].

In the analog domain, data are depicted through continuously varying voltage signals. Various analog-based accelerators have been suggested to execute MAC operations by employing charge manipulation

techniques and Analog-to-Digital Converters (ADCs) [11], [12], [16], [27], [28], [32], [45], [49], [54], [58], [75], [89], [91], [98], [100]. Analog methodologies execute MAC operations within the analog voltage domain utilizing a Static Random Access Memory (SRAM) array, capacitors, and data converters. In these methodologies, input pixel data are encoded either as a Pulse-Width Modulation (PWM) signal or a Pulse-Amplitude-Modulated (PAM) signal. The MAC operation is carried out by summing the read current of simultaneously accessed bit-cells. However, this approach is vulnerable to process variations, noise, bit-flips, and weak line corruption. Despite analog computations demonstrating efficiency in terms of energy (OPS/W) and area (OPS/mm^2), they exhibit limited accuracy and technology scaling due to finite voltage headroom [7].

In Time-Domain (TD) representation, data are depicted as pulses with variable widths or time differences in rising/falling edges, thereby generating variable delays. The TD methodology amalgamates the benefits of both digital and analog approaches; it can scale effectively with technology and offers energy-efficient computation. Furthermore, unlike analog-based computation, which necessitates an analog circuit design flow, TD circuits can employ the digital Integrated Circuit (IC) design flow, facilitating large-scale integration. Prior research indicates that TD cores can outperform digital implementations of Artificial Neural Networks (ANNs) only when the number of input bits is relatively low [2], [7]. In Time-Domain (TD) Artificial Neural Networks (ANNs), calibration becomes necessary due to the analog nature of the delay signal, which is more susceptible to noise and process variation. Additionally, the TD approach necessitates the inclusion of additional components like time-to-digital converters (TDCs) and digital-to-time converters (DTCs). Nonetheless, DTCs and TDCs remain more energy- and area-efficient compared to Digital-to-Analog Converters (DACs) and Analog-to-Digital Converters (ADCs) [63]. TD computing is particularly well-suited for applications requiring low resolution and stringent power constraints, such as edge devices. Phase-Domain (PD) ANN operates similarly to TD, but it employs phase shifts to execute the dot product [90]. The main issues are requiring multiple clock sources and

Table 1.1: Comparing different accelerators

Approach	Digital	Analog	Time
Data representation	Multi-bit digital vector	Continuous voltage signal	Pulse width modulation
Technology scaling	Yes	No	Yes
Immunity to noise	High	Low/moderate	Moderate
Input resolution	High	Low/moderate	Low/moderate
Energy efficiency	Low/moderate	Very high	High
Throughput	High	Very high	Moderate
Area	Moderate	Moderate/large	Moderate/large

the dependence of the toggle activity on the input magnitude. Table 1.1 summarizes the aforementioned approaches.

1.5 Motivation and Scope of the Study

Due to the tremendous number of IoT applications and edge computing where stringent power constraints are required, the need for highly efficient ultra-low-power computing is essential. This work aims to explore and analyze energy-efficient accelerators for edge computing; specifically time-domain and mixed-signal domain cores. Analog computations offer outstanding energy efficiency with real-time parallel processing and learning. This is mainly due to the emerging analog memory technologies which have enabled local storage and processing. Although analog computation is efficient in terms of energy, it has limited accuracy and technology scaling [7]. Additionally, the need for average resolution (e.g., 8 bits) ADCs and DACs limits the efficiency and scalability of the analog cores. Reported works in the literature show that ADCs can contribute up to 64% of total energy consumption [59], which makes it hard to compete against digital accelerators. An emerging trend is to utilize time-domain (TD) to perform MAC operations by representing the data as pulses with modulation, Time-based computation can take advantage of both approaches, analog and digital, as it is energy efficient and can be scaled with CMOS technology.

References

[1] Advanced Technology for High Performance & Low Power Applications, 40 nm Technology—Taiwan Semiconductor Manufacturing Company Limited. URL: https://www.tsmc.com/english/dedicatedFoundry/technology/logic/l_40nm.

[2] H. Al Maharmeh, N. J. Sarhan, C.-C. Hung, M. Ismail, and M. Alhawari, "Compute-in-time for deep neural network accelerators: Challenges and prospects," in *2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 990–993, Springfield, MA, USA, 2020.

[3] H. Al Maharmeh, N. J. Sarhan, C.-C. Hung, M. Ismail, and M. Alhawari, "A comparative analysis of time-domain and digital-domain hardware accelerators for neural networks," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, Daegu, Korea, 2021. DOI: 10.1109/ISCAS51556.2021.9401758.

[4] M. Alhawari, N. Albelooshi, and M. H. Perrott, "A 0.5 V <40 μw, CMOS photoplethysmographic heart-rate sensor IC based on a non-uniform quantizer," in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 384–385, 2013.

[5] F. Alibart, E. Zamanidoost, and D. B. Strukov, "Pattern classification by memristive crossbar circuits using ex situ and in situ training," *Nature Communications*, vol. 4, no. 1, 2013, p. 2072.

[6] B. Alipanahi, A. Delong, M. T. Weirauch, and B. J. Frey, "Predicting the sequence specificities of dna- and rna-binding proteins by deep learning," *Nature Biotechnology*, vol. 33, Jul 2015.

[7] A. Amaravati, S. B. Nasir, J. Ting, I. Yoon, and A. Raychowdhury, "A 55-nm, 1.0–0.4 V, 1.25-pJ/MAC time-domain mixed-signal neuromorphic accelerator with stochastic synapses for reinforcement learning in autonomous mobile robots," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 1, Jan. 2019, pp. 75–87.

[8] S. Ambrogio, P. Narayanan, H. Tsai, R. M. Shelby, I. Boybat, C. di Nolfo, S. Sidler, M. Giordano, M. Bodini, N. C. P. Farinha, B. Killeen, C. Cheng, Y. Jaoudi, and G. W. Burr, "Equivalent-accuracy accelerated neural-network training using analog memory," *Nature*, vol. 558, no. 7708, 2018, pp. 60–67.

[9] A. Andreopoulos, R. Alvarez-Icaza, A. S. Cassidy, and M. D. Flickner, "A low-power neurosynaptic implementation of local binary patterns for texture analysis," in *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 4308–4316, July 2016.

[10] F. M. Bayat, M. Prezioso, B. Chakrabarti, H. Nili, I. Kataeva, and D. Strukov, "Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits," *Nature Communications*, vol. 9, no. 1, 2018, p. 2331.

[11] J. Binas, D. Neil, G. Indiveri, S. Liu, and M. Pfeiffer, "Precise deep neural network computation on imprecise low-power analog hardware," *CoRR,* abs/1606.07786, 2016.

[12] A. Biswas and A. P. Chandrakasan, "Conv-sram: An energy-efficient sram with in-memory dot-product computation for low-power convolutional neural networks," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 1, Jan 2019, pp. 217–230.

[13] G. W. Burr, R. M. Shelby, A. Sebastian, S. Kim, S. Kim, S. Sidler, K. Virwani, M. Ishii, P. Narayanan, A. Fumarola, L. L. Sanches, I. Boybat, M. L. Gallo, K. Moon, J. Woo, H. Hwang, and Y. Leblebici, "Neuromorphic computing using non-volatile memory," *Advances in Physics*, Dec 2017.

[14] G. W. Burr, R. M. Shelby, S. Sidler, C. di Nolfo, J. Jang, I. Boybat, R. S. Shenoy, P. Narayanan, K. Virwani, E. U. Giacometti, B. N. Kurdi, and H. Hwang, "Experimental demonstration and tolerancing of a large-scale neural network (165000 synapses) using phase-change memory as the synaptic weight element," *IEEE Transactions on Electron Devices*, vol. 62, no. 11, Nov 2015, pp. 3498–3507.

[15] F. Cai, J. M. Correll, S. H. Lee, Y. Lim, V. Bothra, Z. Zhang, M. P. Flynn, and W. D. Lu, "A fully integrated reprogrammable memristor-CMOS system for efficient multiply-accumulate operations," *Nature Electronics*, vol. 2, no. 7, 2019, pp. 290–299.

[16] B. Chatterjee, P. Panda, S. Maity, K. Roy, and S. Sen, "An energy-efficient mixed-signal neuron for inherently error-resilient neuromorphic systems," in *Proceedings of the 2017 IEEE International Conference on Rebooting Computing (ICRC)*, pp. 1–2, Nov 2017.

[17] C. Chen, A. Seff, A. Kornhauser, and J. Xiao, "Deepdriving: Learning affordance for direct perception in autonomous driving," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 2722–2730, Dec 2015.

[18] G. K. Chen, R. Kumar, H. E. Sumbul, P. C. Knag, and R. K. Krishnamurthy, "A 4096-neuron 1 m-synapse 3.8-pj/sop spiking neural network with on-chip stdp learning and sparse weights in 10-nm finfet cmos," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 4, April 2019, pp. 992–1002.

[19] Y. Chen, T. Krishna, J. S. Emer, and V. Sze, "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 1, Jan 2017, pp. 127–138.

[20] Z. Chen and J. Gu, "A time-domain computing accelerated image recognition processor with efficient time encoding and non-linear logic operation," *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 54, no. 11, Nov. 2019, pp. 3226–3237.

[21] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," *CoRR,* abs/1409.1259, 2014.

[22] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, Nov. 2011, pp. 2493–2537.

[23] S. Condon, *Facebook unveils Big Basin, new server geared for deep learning, ZDNet*, 2017.

[24] L. Deng, J. Li, J. Huang, K. Yao, D. Yu, F. Seide, M. Seltzer, G. Zweig, X. He, J. Williams, Y. Gong, and A. Acero, "Recent advances in deep learning for speech research at microsoft," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8604–8608, May 2013.

[25] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted mosfet's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, Oct 1974, pp. 256–268.

[26] G. Desoli, N. Chawla, T. Boesch, S. Singh, E. Guidetti, F. De Ambroggi, T. Majo, P. Zambotti, M. Ayodhyawasi, H. Singh, and N. Aggarwal, "14.1 a 2.9 tops/w deep convolutional neural network soc in fd-soi 28 nm for intelligent embedded systems," in *Proceedings of the 2017 IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 238–239, Feb 2017.

[27] Y. Du, L. Du, X. Gu, J. Du, X. S. Wang, B. Hu, M. Jiang, X. Chen, S. S. Iyer, and M. F. Chang, "An analog neural network computing engine using cmos-compatible charge-trap-transistor (ctt)," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2018, pp. 1–1.

[28] M. D. Edwards, H. A. Maharmeh, N. J. Sarhan, M. Ismail, and M. Alhawari, "A low-power, digitally-controlled, multi-stable, CMOS analog memory circuit," in *2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 872–875, Springfield, MA, USA, 2020.

[29] S. K. Esser, P. A. Merolla, J. V. Arthur, A. S. Cassidy, R. Appuswamy, A. Andreopoulos, D. J. Berg, J. L. McKinstry, T. Melano, D. R. Barch, C. di Nolfo, P. Datta, A. Amir, B. Taba, M. D. Flickner, and D. S. Modha, "Convolutional networks for fast, energy-efficient neuromorphic computing," *Proceedings of the National Academy of Sciences*, vol. 113, no. 41, Oct. 2016, pp. 11 441–11 446.

[30] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, Jan 2017, 115 EP–.

[31] L. R. Everson, M. Liu, N. Pande, and C. H. Kim, "An energy-efficient one-shot time-based neural network accelerator employing dynamic threshold error correction in 65 nm," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 10, Oct. 2019, pp. 2777–2785.

[32] L. Fick, D. Blaauw, D. Sylvester, S. Skrzyniarz, M. Parikh, and D. Fick, "Analog in-memory subthreshold deep neural network accelerator," in *Proceedings of the 2017 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1–4, April 2017.

[33] I. Giannopoulos, A. Sebastian, M. Le Gallo, V. P. Jonnalagadda, M. Sousa, M. N. Boon, and E. Eleftheriou, "8-bit precision in-memory multiplication with projected phase-change memory," in *Proceedings of the 2018 IEEE International Electron Devices Meeting (IEDM)*, vol. 4, pp. 27.7.1–27.7, Dec. 2018.

[34] M. Giordano *et al.*, "Analog-to-digital conversion with reconfigurable function mapping for neural networks activation function acceleration," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 2, June 2019, pp. 367–376.

[35] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '14*, pp. 580–587, Washington, DC, USA: IEEE Computer Society, 2014.

[36] T. Gokmen and Y. Vlasov, "Acceleration of deep neural network training with resistive cross-point devices: Design considerations," *Frontiers Neurosci*, vol. 10, Jul. 2016, p. 333.

[37] S. K. Gonugondla, C. Sakr, H. Dbouk, and N. R. Shanbhag, "Fundamental limits on the precision of in-memory architectures," in *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, pp. 1–9, 2020.

[38] A. Gupta *et al.*, "DDPMnet: All-digital pulse density-based DNN architecture with 228 gate equivalents/MAC unit, 28-TOPS/W and 1.5-TOPS/mm2 in 40 nm," in *2022 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1–2, 2022.

[39] R. Z. Han, P. Huang, Y. C. Xiang, C. Liu, Z. Dong, Z. Q. Su, Y. B. Liu, L. Liu, X. Y. Liu, and J. F. Kang, "A novel convolution computing paradigm based on nor flash array with high computing speed and energy efficient," in *Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4, 2018.

[40] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury, and T. Sainath, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Processing Magazine*, vol. 29, November 2012, pp. 82–97.

[41] M. Horowitz, "1.1 computing's energy problem (and what we can do about it)," in *Proceedings of the 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pp. 10–14, Feb 2014.

[42] D. Ielmini and H. S. P. Wong, "In-memory computing with resistive switching devices," *Nature Electronics*, vol. 1, no. 6, 2018, pp. 333–343.

[43] N. P. Jouppi *et al.*, "In-datacenter performance analysis of a tensor processing unit," *CoRR,* abs/1704.04760, 2017.

[44] C.-H. Kim, S. Lim, S. Y. Woo, W.-M. Kang, Y.-T. Seo, S.-T. Lee, S. Lee, D. Kwon, S. Oh, Y. Noh, H. Kim, J. Kim, J.-H. Bae, and J.-H. Lee, "Emerging memory technologies for neuromorphic computing," *Nanotechnology*, vol. 30, no. 3, Nov 2018, p. 032 001.

[45] S. Kim, T. Gokmen, H. Lee, and W. E. Haensch, "Analog cmos-based resistive processing unit for deep neural network training," in *Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 422–425, Aug 2017.

[46] O. Krestinskaya, A. P. James, and L. O. Chua, "Neuro-memristive circuits for edge computing: A review," *CoRR,* abs/1807.00962, 2018.

[47] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.

[48] O. R. N. Laboratory, *Ornl launches summit supercomputer*, 2018.

[49] B. Larras, C. Lahuec, F. Seguin, and M. Arzel, "Ultra-low-energy mixed-signal ic implementing encoded neural networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 11, Nov 2016, pp. 1974–1985.

[50] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE.*, vol. 86, no. 11, Nov. 1998, pp. 2278–2324.

[51] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *Journal of Machine Learning Research*, vol. 17, no. 1, Jan. 2016, pp. 1334–1373.

[52] C. Li, D. Belkin, Y. Li, P. Yan, M. Hu, N. Ge, H. Jiang, E. Montgomery, P. Lin, Z. Wang, W. Song, J. P. Strachan, M. Barnell, Q. Wu, R. S. Williams, J. J. Yang, and Q. Xia, "Efficient and self-adaptive in-situ learning in multilayer memristor neural networks," *Nature Communications*, vol. 9, no. 1, 2018, p. 2385.

[53] C. Li, M. Hu, Y. Li, H. Jiang, N. Ge, E. Montgomery, J. Zhang, W. Song, N. la Davi, C. E. Graves, Z. Li, J. P. Strachan, P. Lin, Z. Wang, M. Barnell, Q. Wu, R. S. Williams, J. J. Yang, and Q. Xia, "Analogue signal and image processing with large memristor crossbars," *Nature Electronics*, vol. 1, no. 1, 2018, pp. 52–59.

[54] Y. Li, S. Kim, X. Sun, P. Solomon, T. Gokmen, H. Tsai, S. Koswatta, Z. Ren, R. Mo, C. C. Yeh, W. Haensch, and E. Leobandung, "Capacitor-based cross-point array for analog neural network with record symmetry and linearity," in *Proceedings of the 2018 IEEE Symposium on VLSI Technology*, pp. 25–26, June 2018.

[55] Z. Li, Y. Chen, L. Gong, L. Liu, D. Sylvester, D. Blaauw, and H. Kim, "An 879 gops 243 mw 80 fps vga fully visual cnn-slam processor for wide-range autonomous exploration," in *Proceedings of the 2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, pp. 134–136, Feb 2019.

[56] M. Liu, L. R. Everson, and C. H. Kim, "A scalable time-based integrate-and-fire neuromorphic core with brain-inspired leak and local lateral inhibition capabilities," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, pp. 1–4, Austin, TX, USA, Apr./May 2017.

[57] H. A. Maharmeh, N. J. Sarhan, M. Ismail, and M. Alhawari, "A 116 TOPS/W spatially unrolled time-domain accelerator utilizing laddered-inverter DTC for energy-efficient edge computing in 65 nm," *IEEE Open Journal of Circuits and Systems*, vol. 4, 2023, pp. 308–323. DOI: 10.1109/OJCAS.2023.3332853.

[58] D. Maliuk and Y. Makris, "An experimentation platform for on-chip integration of analog neural networks: A pathway to trusted and robust analog/rf ics," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 8, 2015, pp. 1721–1734.

[59] M. J. Marinella *et al.*, "Multiscale co-design analysis of energy, latency, area, and accuracy of a ReRAM analog neural training accelerator," *IEEE J. Emerg. Sel. Topics Circuits Syst*, vol. 8, no. 1, Mar. 2018, pp. 86–101.

[60] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, Dec 1943, pp. 115–133.

[61] A. McKenzie, W. Branch, C. Forsythe, and C. D. James, "Toward exascale computing through neuromorphic approaches," Dept. Biosensors, Nanomater., Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. SAND2010-6312, September 2010.

[62] C. Mead, "Neuromorphic electronic systems," *Proceedings of the IEEE*, vol. 78, no. 10, Oct 1990, pp. 1629–1636.

[63] D. Miyashita *et al.*, "An LDPC decoder with time-domain analog and digital mixed-signal processing," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 1, Jan. 2014, pp. 73–83.

[64] D. Miyashita, S. Kousai, T. Suzuki, and J. Deguchi, "A neuromorphic chip optimized for deep learning and CMOS technology with time-domain analog and digital mixed-signal processing," *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 52, no. 10, Oct. 2017, pp. 2679–2689.

[65] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, Feb 2015, 529 EP–.

[66] B. Moons, D. Bankman, L. Yang, B. Murmann, and M. Verhelst, "Binareye: An always-on energy-accuracy-scalable binary cnn processor with all memory on chip in 28 nm cmos," in *Proceedings of the 2018 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1–4, April 2018.

[67] B. Moons and M. Verhelst, "A 0.3–2.6 tops/w precision-scalable processor for real-time largescale convnets," in *Proceedings of the 2016 IEEE Symposium on VLSI Circuits (VLSI-Circuits)*, pp. 1–2, June 2016.

[68] B. Murmann, "Mixed-signal computing for deep neural network inference," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 29, no. 1, Jan. 2021, pp. 3–13.

[69] B. Murmann, D. Bankman, E. Chai, D. Miyashita, and L. Yang, "Mixedsignal circuits for embedded machine-learning applications," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, pp. 1341–1345, Nov. 2015.

[70] Y. Noh, Y. Seo, B. Park, and J. Lee, "Synaptic devices based on 3-d and flash memory architecture for neuromorphic computing," in *Proceedings of the 2019 IEEE 11th International Memory Workshop (IMW)*, pp. 1–4, 2019.

[71] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," *CoRR,* abs/1609.03499, 2016.

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

[73] M. Pfeiffer, M. Schaeuble, J. Nieto, R. Siegwart, and C. Cadena, "From perception to decision: A data-driven approach to end-to-end motion planning for autonomous ground robots," in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1527–1533, May 2017.

[74] M. Price, J. Glass, and A. P. Chandrakasan, "A 6 mw, 5,000-word real-time speech recognizer using wfst models," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 1, Jan 2015, pp. 102–112.

[75] N. Qiao and G. Indiveri, "Analog circuits for mixed-signal neuromorphic computing architectures in 28 nm fd-soi technology," in *Proceedings of the 2017 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, pp. 1–4, Oct 2017.

[76] L. Raczkowski, M. Mozejko, J. Zambonelli, and E. Szczurek, "Ara: Accurate, reliable and active histopathological image classification framework with bayesian deep learning," *Scientific Reports*, vol. 9, no. 1, 2019, p. 14 347.

[77] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, 2015, pp. 211–252.

[78] R. Sarpeshkar, "Analog versus digital: Extrapolating from electronics to neurobiology," *Neural Comput*, vol. 10, no. 7, Oct. 1998, pp. 1601–1638.

[79] A. Sayal, S. S. T. Nibhanupudi, S. Fathima, and J. P. Kulkarni, "A 12.08-TOPS/W all-digital time-domain CNN engine using bi-directional memory delay lines for energy efficient edge computing," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 1, Jan. 2020, pp. 60–75.

[80] A. Sebastian, M. L. Gallo, and E. Eleftheriou, "Computational phase-change memory: Beyond von neumann computing," *Journal of Physics D: Applied Physics*, vol. 52, no. 44, Aug. 2019, p. 443 002.

[81] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "Safe, multi-agent, reinforcement learning for autonomous driving," *CoRR,* abs/1610.03295, 2016.

[82] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, Apr. 2017, pp. 640–651.

[83] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Article Nature*, vol. 529, Jan 2016.

[84] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 1, NIPS'14, pp. 568–576, Cambridge, MA, USA: MIT Press, 2014.

[85] J. Song *et al.*, "TD-SRAM: Time-domain-based in-memory computing macro for binary neural networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 8, Aug. 2021, pp. 3377–3387.

[86] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *CoRR,* abs/1409.3215, 2014.

[87] V. Sze, "Designing hardware for machine learning: The important role played by circuit designers," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 4, Fall 2017, pp. 46–54.

[88] V. Sze, Y. Chen, T. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," *Proceedings of the IEEE*, vol. 105, no. 12, Dec 2017, pp. 2295–2329.

[89] G. Tayfun and Y. Vlasov, "Acceleration of deep neural network training with resistive crosspoint devices," *CoRR,* abs/1603.07341, 2016.

[90] Y. Toyama, K. Yoshioka, K. Ban, S. Maya, A. Sai, and K. Onizuka, "An 8 bit 12.4 tops/w phase-domain mac circuit for energy-constrained deep learning accelerators," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 10, Oct 2019, pp. 2730–2742.

[91] A. Tripathi, M. Arabizadeh, S. Khandelwal, and C. S. Thakur, "Analog neuromorphic system based on multi input floating gate mos neuron model," in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, May 2019.

[92] G. Vision, *CM1K neuromorphic chip with 1024 neurons*, 2017.

[93] E. A. Vittoz, "Future of analog in the VLSI environment," in *Proc. IEEE Int. Symp*, pp. 1372–1375, Circuits Syst, 1990.

[94] C. Wang, W. Lou, L. Gong, L. Jin, L. Tan, Y. Hu, X. Li, and X. Zhou, "Reconfigurable hardware accelerators: Opportunities, trends, and challenges," *CoRR,* abs/1712.04771, 2017.

[95] L. Wang, W. Gao, L. Yu, J.-Z. Wu, and B.-S. Xiong, "Multiple-matrix vector multiplication with crossbar phase-change memory," *Applied Physics Express*, vol. 12, no. 10, 2019, p. 105 002.

[96] R. S. Williams, "What's next? [the end of moore's law]," *Computing in Science Engineering*, vol. 19, no. 2, Mar 2017, pp. 7–13.

[97] Q. Xia and J. J. Yang, "Memristive crossbar arrays for brain-inspired computing," *Nature Materials*, vol. 18, no. 4, 2019, pp. 309–323.

[98] Y. C. Xiang, P. Huang, Z. Zhou, R. Z. Han, Y. N. Jiang, Q. M. Shu, Z. Q. Su, Y. B. Liu, X. Y. Liu, and J. F. Kang, "Analog deep neural network based on nor flash computing array for high speed/energy efficiency computation," in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4, May 2019.

[99] H. Y. Xiong, B. Alipanahi, L. J. Lee, H. Bretschneider, D. Merico, R. K. C. Yuen, Y. Hua, S. Gueroussov, H. S. Najafabadi, T. R. Hughes, Q. Morris, Y. Barash, A. R. Krainer, N. Jojic, S. W. Scherer, B. J. Blencowe, and B. J. Frey, "Rna splicing. the human splicing code reveals new insights into the genetic determinants of disease," *Science (New York, N.Y.)*, vol. 347, no. 6218, Jan 2015, pp. 1 254 806–1 254 806.

[100] M. Yang, C. Yeh, Y. Zhou, J. P. Cerqueira, A. A. Lazar, and M. Seok, "Design of an always-on deep neural network-based 1-mw voice activity detector aided with a customized software model for analog feature extraction," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 6, June 2019, pp. 1764–1777.

[101] S. Yoshimoto, K. Nii, H. Kawaguchi, and M. Yoshimoto, "Multiple-cell-upset hardened 6T SRAM using NMOS-centered layout," in *2013 IEEE International Meeting for Future of Electron Devices*, pp. 98–99, Kansai, 2013.

[102] H. Zeng, M. D. Edwards, G. Liu, and D. K. Gifford, "Convolutional neural network architectures for predicting dna-protein binding," *Bioinformatics (Oxford, England)*, vol. 32, no. 12, Jun 2016, pp. i121–i127.

[103] T. Zhang, G. Kahn, S. Levine, and P. Abbeel, "Learning deep control policies for autonomous aerial vehicles with mpc-guided policy search," in *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 528–535, May 2016.

[104] J. Zhou and O. G. Troyanskaya, "Predicting effects of noncoding variants with deep learning-based sequence model," *Nature Methods*, vol. 12, no. 10, Oct 2015, pp. 931–934.