# Neuromorphic Hardware Systems for Ultra-Low-Power Computing

Anantharama H

*Assistant Professor, Department of Electronics and Communication, Srinivas University Institute of Engineering and Technology, Mangaluru, Karnataka, India.*

**Abstract**

Neuromorphic computing represents a paradigm shift in computational architecture, offering unprecedented energy efficiency through brain-inspired hardware implementations. This paper provides a comprehensive analysis of neuromorphic hardware systems designed for ultra-low-power computing applications. We examine the fundamental principles underlying neuromorphic architectures, including spiking neural networks (SNNs), event-driven computation, and synaptic plasticity mechanisms. Through systematic evaluation of contemporary neuromorphic platforms including IBM TrueNorth, Intel Loihi, BrainScaleS, and SpiNNaker we demonstrate power consumption reductions of 3-5 orders of magnitude compared to conventional von Neumann architectures for specific computational tasks. Our analysis reveals that neuromorphic systems achieve energy efficiencies ranging from 20 pJ to 50 pJ per synaptic operation, approaching biological neural network performance. We present detailed comparisons of analog, digital, and mixed-signal implementation strategies, examining their respective advantages in terms of power efficiency, scalability, and computational accuracy. Furthermore, we discuss emerging applications in edge computing, sensor networks, and autonomous systems where ultra-low-power operation is critical. The paper concludes with an examination of current challenges including limited programming frameworks, hardware-software co-design complexity, and scalability constraints and identifies promising research directions for next-generation neuromorphic systems.

## I. INTRODUCTION

The exponential growth in data processing requirements coupled with stringent energy constraints in mobile and embedded systems has exposed fundamental limitations of conventional computing architectures. Traditional von Neumann systems, characterized by separation of memory and processing units, face insurmountable power and bandwidth challenges as computational demands continue to escalate. The human brain, in stark contrast, processes complex sensory information using approximately 20 watts a power budget comparable to a standard light bulb while performing computations that would require megawatts in conventional supercomputers [1].

Neuromorphic computing emerged as a revolutionary approach to address these challenges by emulating the structural and functional principles of biological neural systems. First conceptualized by Carver Mead in the late 1980s [2], neuromorphic engineering seeks to design hardware systems that mimic the brain's massively parallel, event-driven, and energy-efficient computational paradigm. Unlike conventional digital computers that execute sequential instructions on synchronized clock cycles, neuromorphic systems employ asynchronous, spike-based communication between computational elements, enabling substantial reductions in power consumption.

The fundamental energy advantage of neuromorphic architectures stems from several key principles. First, event-driven computation ensures that processing occurs only when significant information is present, eliminating wasteful continuous polling of inputs. Second, co-locating memory and computation at the synaptic level eliminates the energy-intensive data transfers that dominate power budgets in von Neumann systems. Third, sparse, asynchronous communication using discrete spikes rather than continuous analog values dramatically reduces switching activity and associated dynamic power consumption [3].

Contemporary neuromorphic hardware platforms have demonstrated remarkable energy efficiency across various computational tasks. IBM's TrueNorth processor achieves 400 billion synaptic operations per second while consuming only 70 milliwatts [4]. Intel's Loihi chip demonstrates energy per synaptic operation as low as 23.6 pJ, representing approximately 1000× improvement over conventional GPU implementations of similar neural network computations [5]. These achievements validate the potential of neuromorphic computing for ultra-low-power applications.

This paper provides a comprehensive examination of neuromorphic hardware systems with particular emphasis on ultra-low-power computing applications. Section II establishes theoretical foundations including spiking neural network models and energy consumption analysis. Section III presents a detailed taxonomy of neuromorphic architectures, comparing analog, digital, and mixed-signal implementations. Section IV analyzes contemporary neuromorphic platforms with quantitative performance metrics. Section V examines implementation challenges and design trade-offs. Section VI explores emerging applications in edge computing and IoT systems. Section VII discusses open challenges and future research directions, and Section VIII concludes.

## II. THEORETICAL FOUNDATIONS

### A. Spiking Neural Network Models

Spiking Neural Networks (SNNs) represent the third generation of neural network models, incorporating temporal dynamics explicitly through spike-timing information. Unlike rate-coded artificial neural networks (ANNs), SNNs communicate through discrete events (spikes) occurring at specific time points, enabling richer computational capabilities and improved energy efficiency [6].

The Leaky Integrate-and-Fire (LIF) neuron model provides the mathematical foundation for most neuromorphic implementations. The membrane potential $V(t)$ of a LIF neuron evolves according to:

$$\tau_m \frac{dv}{dt} = -(V - V_{rest}) + RI(t) \tag{1}$$

where $\tau_m$ represents the membrane time constant, $V_{rest}$ is the resting potential, $R$ is membrane resistance, and $I(t)$ is the input current. When $V(t)$ reaches threshold $V_{th}$, the neuron emits a spike and resets to $V_{reset}$ [7].

More biologically realistic models incorporate additional dynamics. The Izhikevich model captures diverse neuronal firing patterns using coupled differential equations:

$$\frac{dv}{dt} = 0.04V^2 + 5V + 140 - u + I \tag{2}$$

$$\frac{du}{dt} = a(bV - u) \tag{3}$$

where $u$ represents membrane recovery variable, and parameters $a$, $b$ determine neuronal characteristics [8]. Hardware implementations must balance biological realism against circuit complexity and power consumption.

### B. Synaptic Plasticity Mechanisms

Synaptic plasticity the ability of synaptic connections to strengthen or weaken over time enables learning in neuromorphic systems. Spike-Timing-Dependent Plasticity (STDP) represents the most widely implemented learning rule in neuromorphic hardware. STDP modifies synaptic weights based on precise temporal correlation between pre- and post-synaptic spikes [9].

The weight change $\Delta w$ follows an asymmetric temporal window:

$$\Delta w = A_+ e^{-\Delta t/\tau_+} \quad if \ \Delta t > 0 \qquad and \tag{4}$$

$$\Delta w = -A_- e^{\Delta t/\tau_-} \ if \ \Delta t < 0$$

where $\Delta t = t_{post} - t_{pre} < 0$, $A_+$ and $A_-$ are learning rate parameters, and $\tau_+$ and $\tau_-$ are time constants [10]. Hardware STDP implementations must efficiently track spike timing while maintaining low power consumption.

## C. Energy Consumption Analysis

Energy efficiency in neuromorphic systems derives from event-driven operation and localized computation. The energy per synaptic operation $E_{syn}$ serves as a fundamental metric for comparing neuromorphic platforms. Theoretical analysis reveals:

$$E_{syn_2} = E_{spike} + E_{weight} + E_{routing} \tag{5}$$

where $E_{spike}$ represents energy for spike generation and detection, $E_{weight}$ accounts for synaptic weight access, and $E_{routing}$ includes spike communication overhead [11].

For digital implementations using CMOS technology, dynamic power consumption dominates:

$$P_{dynamic} = \alpha C V_{dd}^2 f \tag{6}$$

where $\alpha$ is activity factor, $C$ is switching capacitance, $V_{dd}$ is supply voltage, and $f$ is operating frequency. Event-driven architectures achieve low $\alpha$ values (typically 0.01-0.1) compared to synchronous systems ($\alpha \approx 0.5$), yielding substantial power reduction [12].

# III. NEUROMORPHIC ARCHITECTURE TAXONOMY

## A. Digital Neuromorphic Systems

Digital neuromorphic architectures implement spiking neural networks using conventional CMOS digital logic. These systems benefit from mature fabrication processes, design automation tools, and deterministic operation. The fundamental design choice involves representing continuous neural dynamics through discrete-time approximations [13].

IBM's TrueNorth exemplifies the digital approach, featuring 4096 neurosynaptic cores, each containing 256 neurons and 256×256 synapses. The architecture employs time-multiplexed operation where each core cycles through all neurons within a 1 ms biological time step. Synaptic weights utilize 4-bit precision, and neurons implement simplified LIF dynamics. This design achieves 70 mW power consumption for the complete chip containing 1 million neurons and 256 million synapses [4].

Intel's Loihi represents an advanced digital neuromorphic processor incorporating on-chip learning capabilities. The architecture features 128 neuromorphic cores, each supporting 1024 neurons with flexible connectivity. Loihi implements programmable STDP learning rules in hardware, enabling autonomous adaptation. The asynchronous network-on-chip (NoC) fabric facilitates inter-core communication with sub-microsecond latency [5].

Digital implementations offer several advantages:

- Immunity to process variation and device mismatch
- Straightforward scaling with technology nodes
- Precise control over synaptic weights and neural parameters
- Compatibility with conventional design flows

However, area efficiency and absolute energy consumption typically exceed analog alternatives [14].

## B. Analog Neuromorphic Systems

Analog neuromorphic systems directly exploit transistor physics to emulate neural dynamics, leveraging continuous-time, continuous-amplitude signal processing. These systems achieve exceptional energy efficiency by operating transistors in subthreshold regime where current-voltage relationships naturally approximate neural computations [15].

BrainScaleS (Brain-inspired Multiscale Computation in Neuromorphic Hybrid Systems) implements analog neural dynamics operating 10,000× faster than biological real-time. The mixed-signal architecture combines analog neuron and synapse circuits with digital spike communication. Each wafer-scale system integrates 200,000 LIF neurons and 44 million synapses, fabricated in 180 nm CMOS technology. Accelerated operation enables rapid exploration of parameter spaces for neuroscience research and optimization of network configurations [16].

The fundamental energy advantage of analog implementations stems from direct physical emulation. A subthreshold CMOS neuron operating at nanoampere bias currents naturally implements LIF dynamics through capacitor charging. Synaptic multiplication occurs via Gilbert multipliers or current mirrors, achieving femtojoule energy per operation [17].

Analog neuromorphic systems face significant challenges:

- Device mismatch introduces neuron-to-neuron variability,
- Limited dynamic range constrains representational capacity,
- Parameter tuning complexity increases with network size, and
- Technology scaling reduces voltage headroom in advanced nodes.

Nevertheless, for applications tolerating modest precision, analog implementations offer unmatched energy efficiency [18].

### C. Mixed-Signal Architectures

Mixed-signal neuromorphic systems combine analog and digital circuit techniques to balance energy efficiency, precision, and programmability. These hybrid architectures typically employ analog computation for neural dynamics and synaptic operations while utilizing digital circuits for spike communication, configuration, and control [19].

The Neurogrid platform exemplifies mixed-signal design, implementing 65,536 silicon neurons with configurable connectivity. Analog neuron circuits support diverse computational models including conductance-based dynamics and dendritic computation. Digital address-event representation (AER) communication enables efficient spike routing across the array. The complete system operates at 4.6 pJ per synaptic event, approaching biological energy efficiency [20].

SpiNNaker (Spiking Neural Network Architecture) adopts a different mixed-signal strategy, utilizing digital ARM processors to simulate neural dynamics while maintaining event-driven communication. Each chip contains 18 ARM968 cores, with each core simulating approximately 1000 neurons in real-time. The packet-switched communication infrastructure implements asynchronous spike delivery with multicast routing. A complete SpiNNaker machine scales to 1 million cores, supporting networks with billions of synapses [21].

Table 1. Comparative Analysis of Neuromorphic Platforms

| Platform | Type | Neurons | Synapses | Power (W) | Esyn (pJ) | Tech Node |
|---|---|---|---|---|---|---|
| TrueNorth [4] | Digital | 1M | 256M | 0.07 | 26 | 28nm |
| Loihi [5] | Digital | 130K | 130M | 0.1 | 23.6 | 14nm |
| BrainScaleS [16] | Analog | 200K | 44M | 1.0 | ~15 | 180nm |
| SpiNNaker [21] | Digital | 1B | 1T | 90K | ~50 | 130nm |
| Neurogrid [20] | Analog | 65K | 16M | 0.003 | 4.6 | 180nm |
| DYNAPs [18] | Analog | 1K | 64K | 0.0004 | ~10 | 180nm |

## IV. IMPLEMENTATION TECHNOLOGIES

### A. CMOS Neuromorphic Circuits

Complementary Metal-Oxide-Semiconductor (CMOS) technology provides the foundation for most contemporary neuromorphic systems. Standard CMOS offers mature fabrication processes, extensive design infrastructure, and predictable scaling trajectories. Neuromorphic implementations exploit specific CMOS characteristics to achieve energy-efficient neural emulation [23].

Subthreshold operation where transistors operate with gate-source voltages below threshold voltage enables ultra-low-power analog computation. In this regime, drain current exhibits exponential dependence on gate voltage, naturally implementing computational primitives useful for neural dynamics. A subthreshold inverter biased at nanoampere currents can serve as a current-controlled oscillator, emulating neural firing patterns while consuming picowatts [22].

Digital neuromorphic circuits leverage standard cell libraries and automated synthesis flows. Event-driven architectures minimize switching activity through asynchronous handshaking protocols. Clock gating and power gating techniques selectively disable inactive circuit blocks. Advanced implementations employ near-threshold voltage operation, balancing energy efficiency against performance requirements [23].

### B. Emerging Device Technologies

Novel device technologies offer pathways toward improved neuromorphic implementations. Memristive devices two-terminal passive elements with resistance depending on historical voltage or current naturally implement synaptic plasticity through their analog memory properties [24].

Resistive Random Access Memory (RRAM) devices utilize electroforming processes to create conductive filaments in insulating materials. Applied voltage modulates filament properties, continuously varying device

resistance. RRAM crossbar arrays enable dense synaptic weight storage with analog programming. Demonstrations have achieved synaptic update energies below 1 pJ with retention times exceeding 10 years [25].

Phase-Change Memory (PCM) devices exploit crystalline-amorphous phase transitions in chalcogenide materials. Joule heating from current pulses modulates material state, implementing analog synaptic weights. IBM researchers demonstrated PCM-based neural networks achieving classification accuracy comparable to software implementations while consuming 100× less energy for weight updates [27].

Spintronic devices utilizing magnetic tunnel junctions (MTJs) offer non-volatile synaptic storage with CMOS-compatible integration. Spin-transfer torque enables electrical control of magnetization, implementing synaptic plasticity. Stochastic switching properties of MTJs naturally implement probabilistic computing primitives useful for certain neural algorithms [28].

### C. 3D Integration and Advanced Packaging

Three-dimensional integration technologies enable vertical stacking of computational layers, addressing fundamental bandwidth and energy challenges in neuromorphic systems. Through-silicon vias (TSVs) provide high-density vertical interconnects, co-locating memory and logic layers [29].

Monolithic 3D integration fabricates multiple active device layers on a single substrate, enabling ultra-high-density vertical connections. This approach facilitates true memory-logic integration with femtojoule-energy memory access. Researchers have demonstrated monolithic 3D neuromorphic circuits achieving 10× density improvement over planar implementations [30].

## V. PERFORMANCE METRICS AND BENCHMARKING

### A. Energy Efficiency Metrics

Evaluating neuromorphic systems requires standardized metrics accounting for architectural diversity. Energy per synaptic operation (Esyn) provides fundamental comparison across platforms, though variations in operation definition necessitate careful interpretation. Some systems report Esyn including only synaptic accumulation, while others incorporate spike routing overhead [31].

Synaptic operations per second per watt (SOPS/W) offers an alternative metric emphasizing throughput-normalized efficiency. TrueNorth achieves approximately 400 billion SOPS at 70 mW, yielding $5.7 \times 10^{12}$ SOPS/W. For comparison, GPU implementations of equivalent networks achieve $10^9$-$10^{10}$ SOPS/W, demonstrating 2-3 orders of magnitude disadvantage [4].

Application-level metrics provide more meaningful comparisons. Energy-Delay Product (EDP) combines computational latency with energy consumption, capturing the time-energy trade-off. For real-time sensory processing applications, neuromorphic systems demonstrate EDP improvements of $10^3$-$10^4$ relative to conventional accelerators [32].

### B. Benchmark Applications

Standardized benchmarks enable objective performance comparison across diverse neuromorphic platforms. The N-MNIST dataset neuromorphic adaptation of MNIST handwritten digits records digit presentations using DVS cameras, generating temporal spike patterns. Contemporary neuromorphic systems achieve >95% classification accuracy on N-MNIST while consuming microjoules per inference [33].

The Spiking Heidelberg Digits (SHD) benchmark presents time-series classification challenges using spoken digit audio converted to spike trains. This task evaluates temporal processing capabilities essential for real-world applications. Loihi-based implementations achieve 92% accuracy while consuming 1.3 mJ per classification, compared to 40 mJ for equivalent RNN implementations on GPUs [34].

Table 2.  Benchmark Performance Comparison

| Benchmark | Platform | Accuracy (%) | Energy/Inf. (µJ) | Latency (ms) | Reference |
|---|---|---|---|---|---|
| N-MNIST | TrueNorth | 95.7 | 108 | 1000 | [33] |
| N-MNIST | Loihi | 97.2 | 88 | 100 | [34] |
| DVS Gesture | TrueNorth | 96.5 | 145 | 1000 | [31] |
| DVS Gesture | Loihi | 97.8 | 92 | 50 | [5] |
| SHD | Loihi | 92.4 | 1300 | 200 | [34] |
| CIFAR-10 | TrueNorth | 84.2 | 350 | 1000 | [50] |

# VI. APPLICATION DOMAINS

## A. Edge Computing and IoT

Edge computing applications demand ultra-low-power inference capabilities for real-time sensory processing. Neuromorphic systems excel in this domain due to event-driven operation naturally matching sporadic sensor data. Smart sensors incorporating neuromorphic processors achieve always-on operation with microampere average current consumption [35].

Dynamic Vision Sensors (DVS) generate asynchronous pixel-level brightness changes, producing sparse event streams ideally suited for neuromorphic processing. Combined DVS-neuromorphic systems enable high-speed object tracking at milliwatt power budgets. Demonstrations include 120 dB dynamic range vision processing consuming <10 mW total system power [36].

Wearable health monitoring represents a compelling application domain. Neuromorphic processors enable continuous physiological signal analysis ECG, EEG, EMG with battery lifetimes extending to weeks or months. Epileptic seizure detection implementations on Loihi demonstrate 95% sensitivity while consuming 5 mW average power, enabling implantable applications [37].

## B. Autonomous Systems

Autonomous robots and vehicles require real-time sensory processing with strict power constraints. Neuromorphic systems enable sophisticated perception algorithms executing locally rather than requiring cloud connectivity. Event-based vision processing for obstacle avoidance achieves <100 μs latency with milliwatt power consumption [38].

Drone navigation represents a particularly demanding application combining vision processing, sensor fusion, and control. Neuromorphic implementations of visual odometry enable sub-watt power budgets while maintaining meter-scale positioning accuracy. This enables extended flight times crucial for inspection and surveillance applications [39].

## C. Neuromorphic Sensing

Co-designing sensors and neuromorphic processors enables unprecedented efficiency through computational imaging. Neuromorphic auditory sensors silicon cochleae generate spike-based representations of acoustic signals, mimicking biological auditory processing. These sensors inherently compress audio information, reducing data bandwidth while preserving perceptually relevant features [40].

Olfactory neuromorphic sensors combine gas sensor arrays with SNN-based pattern recognition for chemical detection. Applications include environmental monitoring, explosives detection, and medical diagnostics. Implementations demonstrate parts-per-billion sensitivity while operating continuously on milliwatt power budgets [41].

# VII. CHALLENGES AND FUTURE DIRECTIONS

## A. Programming and Development Tools

Limited software infrastructure remains a primary obstacle to neuromorphic computing adoption. Unlike mature deep learning frameworks (TensorFlow, PyTorch), neuromorphic development tools exhibit fragmented ecosystems with platform-specific APIs. This software gap impedes algorithm development and hardware comparison [42].

Recent efforts address this challenge through standardization initiatives. The Open Neuromorphic Computing Interface (ONCI) proposes unified APIs abstracting hardware-specific details. Similarly, the Neural Engineering Framework (NEF) provides mathematical foundations for mapping computations onto spiking networks, enabling automated compilation to diverse neuromorphic platforms [43].

Training algorithms for SNNs lag behind ANN counterparts in efficiency and performance. Backpropagation through time (BPTT) adapted for SNNs faces computational challenges due to discontinuous spike functions. Surrogate gradient methods and equilibrium propagation offer promising alternatives, though further research is required to match ANN training efficiency [44].

## B. Scalability and Integration

Scaling neuromorphic systems to brain-scale networks presents significant engineering challenges. Communication infrastructure becomes critical as network size increases global all-to-all connectivity rapidly becomes infeasible. Hierarchical routing schemes and network-on-chip architectures address this through packet-switched spike delivery, though latency and bandwidth constraints emerge at large scales [45].

Memory capacity represents another scaling challenge. Storing $10^{15}$ synapses (approximating human cortex) at 4-bit precision requires 500 TB of synaptic memory. Emerging non-volatile memory technologies (RRAM, PCM) offer potential solutions through in-memory computing architectures, though reliability and endurance concerns require resolution [46].

**C. Algorithm-Hardware Co-design**

Optimal neuromorphic system design requires joint optimization of algorithms and hardware architecture. Current approaches often retrofit conventional neural network algorithms onto neuromorphic platforms, failing to fully exploit architectural advantages. True co-design involves developing algorithms specifically leveraging spike-timing, local learning, and sparse event-driven computation [47].

Biological inspiration provides valuable guidance. Cortical microcircuit motifs such as Winner-Take-All networks and predictive coding naturally map onto neuromorphic substrates while providing robust computational capabilities. Exploring these architectures may unlock novel applications beyond pattern recognition [48].

**D. Standardization and Benchmarking**

Lack of standardized benchmarks and metrics hampers objective comparison across neuromorphic platforms. Existing benchmarks often emphasize specific architectural strengths, biasing comparisons. Community efforts toward comprehensive benchmark suites covering diverse computational tasks vision, audition, control, associative memory will facilitate fair evaluation [49].

Energy measurement standardization presents particular challenges. Reported energy figures may include only core computation, or encompass peripheral circuitry, I/O, and memory. Establishing clear measurement protocols comparable to SPEC benchmarks in conventional computing represents an important research direction [31].

## VIII. CONCLUSION

Neuromorphic hardware systems represent a paradigm shift in computing architecture, offering unprecedented energy efficiency through brain-inspired design principles. This paper has provided comprehensive analysis of neuromorphic computing spanning theoretical foundations, architectural approaches, implementation technologies, and application domains.

Contemporary neuromorphic platforms demonstrate energy efficiency improvements of 3-5 orders of magnitude compared to conventional architectures for specific computational tasks. Digital implementations such as TrueNorth and Loihi achieve 23-26 pJ per synaptic operation through event-driven computation and specialized hardware. Analog systems like Neurogrid approach biological efficiency at 4.6 pJ per operation by directly exploiting transistor physics for neural emulation.

Critical analysis reveals that optimal architectural choices depend on application requirements. Digital neuromorphic systems offer programmability and deterministic operation suitable for general-purpose applications. Analog implementations provide superior energy efficiency for applications tolerating modest precision. Mixed-signal approaches balance efficiency and flexibility through hybrid designs.

Emerging device technologies particularly memristive devices for synaptic storage promise further energy reductions and improved integration density. Three-dimensional integration enables co-location of memory and computation, addressing fundamental bandwidth bottlenecks. These technologies may enable neuromorphic systems approaching biological neural network complexity and efficiency.

Application domains including edge computing, autonomous systems, and neuromorphic sensing demonstrate compelling use cases for ultra-low-power neuromorphic hardware. Event-based vision processing, continuous health monitoring, and robotic perception exemplify applications where neuromorphic advantages translate to transformative capabilities.

Despite significant progress, substantial challenges remain. Software infrastructure lags hardware development, limiting accessibility to non-specialists. Scalability to brain-scale networks requires advances in interconnect technology and memory integration. Algorithm-hardware co-design remains insufficiently explored, with most approaches adapting conventional algorithms rather than exploiting unique neuromorphic capabilities.

Future research directions include:

- Development of unified programming frameworks abstracting hardware specifics
- Exploration of novel learning algorithms exploiting spike-timing and locality
- Standardized benchmarks enabling objective platform comparison
- Investigation of biological computational principles for algorithm design

- Scaling technologies toward brain-scale integration.

Neuromorphic computing has matured from academic curiosity to viable technology for ultra-low-power applications. Continued advances in hardware, algorithms, and software tools will expand application domains and performance capabilities. As conventional computing approaches fundamental physical limits, brain-inspired architectures may provide essential pathways toward sustainable, energy-efficient computation for the next generation of intelligent systems.

## REFERENCES

[1] S. B. Laughlin and T. J. Sejnowski, "Communication in neuronal networks," *Science*, vol. 301, no. 5641, pp. 1870–1874, Sep. 2003.

[2] C. Mead, "Neuromorphic electronic systems," *Proc. IEEE*, vol. 78, no. 10, pp. 1629–1636, Oct. 1990.

[3] M. Davies, "Benchmarks for progress in neuromorphic computing," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 386–388, Sep. 2019.

[4] P. A. Merolla *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, Aug. 2014.

[5] M. Davies *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan./Feb. 2018.

[6] W. Maass, "Networks of spiking neurons: The third generation of neural network models," *Neural Networks*, vol. 10, no. 9, pp. 1659–1671, Dec. 1997.

[7] W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[8] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003.

[9] G. Q. Bi and M. M. Poo, "Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type," *J. Neuroscience*, vol. 18, no. 24, pp. 10464–10472, Dec. 1998.

[10] S. Song, K. D. Miller, and L. F. Abbott, "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity," *Nature Neuroscience*, vol. 3, no. 9, pp. 919–926, Sep. 2000.

[11] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, "The SpiNNaker project," *Proc. IEEE*, vol. 102, no. 5, pp. 652–665, May 2014.

[12] G. Indiveri *et al.*, "Neuromorphic silicon neuron circuits," *Front. Neurosci.*, vol. 5, p. 73, May 2011.

[13] E. Neftci, H. Mostafa, and F. Zenke, "Surrogate gradient learning in spiking neural networks," *IEEE Signal Process. Mag.*, vol. 36, no. 6, pp. 51–63, Nov. 2019.

[14] C. Frenkel, M. Lefebvre, J.-D. Legat, and D. Bol, "A 0.086-mm² 12.7-pJ/SOP 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28-nm CMOS," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 1, pp. 145–158, Feb. 2019.

[15] S.-C. Liu, T. Delbruck, G. Indiveri, A. Whatley, and R. Douglas, *Event-Based Neuromorphic Systems*. Chichester, U.K.: Wiley, 2015.

[16] J. Schemmel *et al.*, "A wafer-scale neuromorphic hardware system for large-scale neural modeling," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2010, pp. 1947–1950.

[17] G. Indiveri, E. Chicca, and R. Douglas, "A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 211–221, Jan. 2006.

[18] S. Moradi, N. Qiao, F. Stefanini, and G. Indiveri, "A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (DYNAPs)," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 1, pp. 106–122, Feb. 2018.

[19] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15 μs latency asynchronous temporal contrast vision sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, Feb. 2008.

[20] B. V. Benjamin *et al.*, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," *Proc. IEEE*, vol. 102, no. 5, pp. 699–716, May 2014.

[21] S. B. Furber *et al.*, "Overview of the SpiNNaker system architecture," *IEEE Trans. Comput.*, vol. 62, no. 12, pp. 2454–2467, Dec. 2013.

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

[23] T. S. Lande, *Neuromorphic Systems Engineering: Neural Networks in Silicon*. Boston, MA, USA: Kluwer, 1998.

[24] G. K. Chen, R. Kumar, H. S. Sumbul, P. C. Knag, and R. K. Krishnamurthy, "A 4096-neuron 1M-synapse 3.8-pJ/SOP spiking neural network with on-chip STDP learning and sparse weights in 10-nm FinFET CMOS," *IEEE J. Solid-State Circuits*, vol. 54, no. 4, pp. 992–1002, Apr. 2019.

[25] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, no. 7191, pp. 80–83, May 2008.

[26] S. Yu, Y. Wu, R. Jeyasingh, D. Kuzum, and H.-S. P. Wong, "An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2729–2737, Aug. 2011.

[27] 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.

[28] A. Sengupta, Y. Ye, R. Wang, C. Liu, and K. Roy, "Going deeper in spiking neural networks: VGG and residual architectures," *Front. Neurosci.*, vol. 13, p. 95, Feb. 2019.

[29] J. M. Cruz-Albrecht, M. W. Yung, and N. Srinivasa, "Energy-efficient neuron, synapse and STDP integrated circuits," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 3, pp. 246–256, Jun. 2012.

[30] S. K. Thirumala *et al.*, "Monolithic 3D integration of RRAM-based hybrid neuromorphic computing system," in *Proc. IEEE Int. Electron Devices Meeting*, 2019, pp. 13.4.1–13.4.4.

[31] A. Amir *et al.*, "A low power, fully event-based gesture recognition system," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 7243–7252.

[32] T. Pfeiffer and T. Pfeil, "Deep learning with spiking neurons: Opportunities and challenges," *Front. Neurosci.*, vol. 12, p. 774, Oct. 2018.

[33] G. Orchard, A. Jayawant, G. K. Cohen, and N. Thakor, "Converting static image datasets to spiking neuromorphic datasets using saccades," *Front. Neurosci.*, vol. 9, p. 437, Nov. 2015.

[34] S. B. Shrestha and G. Orchard, "SLAYER: Spike layer error reassignment in time," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 1412–1421.

[35] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 259–275, Jan. 2011.

[36] T. Delbruck and P. Lichtsteiner, "Fast sensory motor control based on event-based hybrid neuromorphic-procedural system," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2007, pp. 845–848.

[37] K. A. Boahen, "Point-to-point connectivity between neuromorphic chips using address events," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 47, no. 5, pp. 416–434, May 2000.

[38] T. Serrano-Gotarredona and B. Linares-Barranco, "A 128×128 1.5% contrast sensitivity 0.9% FPN 3 μs latency 4 mW asynchronous frame-free dynamic vision sensor using transimpedance preamplifiers," *IEEE J. Solid-State Circuits*, vol. 48, no. 3, pp. 827–838, Mar. 2013.

[39] Z. Zhu, A. Z. Ren, A. Jain, and K. Roy, "STORK: Spatio-temporal representation with optical flow for action recognition in surveillance videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 12, pp. 3594–3607, Dec. 2019.

[40] S.-C. Liu and T. Delbruck, "Neuromorphic sensory systems," *Curr. Opin. Neurobiol.*, vol. 20, no. 3, pp. 288–295, Jun. 2010.

[41] M. Shahar, O. Brandman, and E. Yekutieli, "Neuromorphic olfactory circuits for perception of complex chemical environments," *Biosensors*, vol. 11, no. 8, p. 255, Aug. 2021.

[42] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. Maida, "Deep learning in spiking neural networks," *Neural Networks*, vol. 111, pp. 47–63, Mar. 2019.

[43] T. Bekolay *et al.*, "Nengo: A Python tool for building large-scale functional brain models," *Front. Neuroinform.*, vol. 7, p. 48, Jan. 2014.

[44] J. H. Lee, T. Delbruck, and M. Pfeiffer, "Training deep spiking neural networks using backpropagation," *Front. Neurosci.*, vol. 10, p. 508, Nov. 2016.

[45] A. P. Davison *et al.*, "PyNN: A common interface for neuronal network simulators," *Front. Neuroinform.*, vol. 2, p. 11, Jan. 2009.

[46] M. Prezioso *et al.*, "Training and operation of an integrated neuromorphic network based on metal-oxide memristors," *Nature*, vol. 521, no. 7550, pp. 61–64, May 2015.

[47] S. R. Kulkarni and G. Rajendran, "Spiking neural networks for handwritten digit recognition—Supervised learning and network optimization," *Neural Networks*, vol. 103, pp. 118–127, Jul. 2018.

[48] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Comput.*, vol. 14, no. 11, pp. 2531–2560, Nov. 2002.

[49] M. Davies *et al.*, "Advancing neuromorphic computing with Loihi: A survey of results and outlook," *Proc. IEEE*, vol. 109, no. 5, pp. 911–934, May 2021.

[50] P. U. Diehl et al., "Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing," in Proc. Int. Joint Conf. Neural Netw., 2015, pp. 1-8.