

Neuromorphic Chips: Scalable Memristor Crossbar Architectures for Energy-Efficient Edge-AI

Nanda Kishor S¹, Nithin Joe²

¹UG Scholar, Department of ECE, NCERC, Thrissur, Kerala

²Assistant professor, Department of ECE, NCERC, Thrissur, Kerala

Abstract - Neuromorphic computing has emerged as a revolutionary area in contemporary information processing by emulating the architecture and operation of the human brain to enable artificial intelligence. In contrast with traditional von Neumann systems like CPUs and GPUs, which are plagued by excessive energy use, latency, and limited scalability because of the divide between memory and computation, neuromorphic chips integrate these two functions into a single platform. At the center of this architecture are memristor crossbar arrays, which support in-memory computation, parallel processing, and adaptive learning, thus eliminating the von Neumann bottleneck. Such integration is especially critical for Edge-AI applications, where devices are required to provide real-time intelligence, work with tight energy budgets, and work without relying on cloud resources. Based on the seminar theme, this review focuses on scalable neuromorphic chip architectures and their uses in applications ranging from robotics, healthcare, IoT, and smart cities. Comparative evaluation with traditional processors features enhanced energy efficiency, computational throughput, and on-chip adaptability. However, issues like memristor variability, device endurance, and fabrication complexity are still open issues for large-scale deployment. Neuromorphic chips stand at the intersection of potential future research avenues, ranging from hybrid CMOS–memristor integration to three-dimensional crossbar structures and algorithm–hardware co-design, as a central enabler of next-generation, energy-efficient, intelligent edge computing systems.

Key Words: neuromorphic chip, memristor crossbar, Edge-AI, in-memory computing, spiking neural networks, low power.

1. INTRODUCTION

Neuromorphic computing has emerged as a revolutionary discipline in contemporary information processing by modeling the structure and operations of the human brain to be used in artificial intelligence. Neuromorphic chips combine the functions of computation and memory into a single platform, unlike traditional von Neumann architectures like CPUs and GPUs that are plagued by high power consumption, latency, and low scalability owing to the dissociation of memory and computation. At the center of this design are memristor-based crossbar arrays, which allow in-memory computation, parallel processing, and adaptive learning and thus provide a solution to the von Neumann bottleneck. This convergence is especially vital for Edge-AI applications, where devices need to provide real-time intelligence, work with very tight energy budgets, and

work autonomously from the cloud. Based on the seminar theme, this review focuses on scalable neuromorphic chip architectures and their use in applications like robotics, healthcare, IoT, and smart cities. Comparative study with traditional processors underscores gains in energy efficiency, computation throughput, and on-chip adaptability. Concurrently, issues like memristor variability, device endurance, and fabrication complexity are still open questions to large-scale deployment. Future research areas such as hybrid CMOS–memristor integration, three-dimensional crossbar arrays, and algorithm–hardware co-design place neuromorphic chips at the center of next-generation, energy-efficient, and intelligent edge computing systems. Neuromorphic computing is a paradigm shift in information processing for the present era as it directly takes inspiration from the architecture and functionality of the human brain. Traditional computing systems, which are based on the von Neumann model, have distinct memory and processing units. This isolation presents a significant performance bottleneck referred to as the "von Neumann bottleneck," where repeated data movement between processor and memory results in high latency, high energy consumption, and poor scalability. These have become particularly important with the advent of Artificial Intelligence (AI), when low-latency decision-making and energy-efficient computation are necessary for edge devices in robotics, healthcare, and the Internet of Things (IoT).

Neuromorphic chips resolve these constraints by integrating storage and computation onto a shared platform. At their core are memristor-based crossbar arrays, which are artificial synapses and support in-memory computing, parallel matrix–vector multiplications, and adaptive learning via mechanisms like spike-timing dependent plasticity (STDP). This neuroinspired strategy enables neuromorphic systems to operate at ultra-low power consumption and in real time, which makes them extremely well-suited for Edge-AI deployments where cloud connectivity is constrained or undesirable.

Recent studies and prototype advancements show that neuromorphic chips not just enhance computation throughput but also enable on-chip learning and fault tolerance with dynamic workloads. Applications emphasized by seminar presentations are autonomous navigation, biomedical signal processing, smart city

infrastructure, and low-power embedded AI systems. Meanwhile, the area is still challenged by memristor variability, endurance constraints, fabrication complexity, and integration into CMOS technology.

Considering the intense growth of neuromorphic research, a scientific exploration of scalable chip design is timely and warranted. This paper integrates findings from the seminar and literature to assess neuromorphic chip design, contrasts them with conventional processors, and investigates their appropriateness for practical Edge-AI systems. The review also establishes open problems and areas for research—such as hybrid CMOS–memristor integration, 3D crossbar architectures, and algorithm–hardware co-design—planning neuromorphic chips as a solid foundation for the next generation of intelligent and energy-efficient edge computing.

2. LITERATURE REVIEW

Xiong et al. [1] have provided a review of neuromorphic computing models and hardware platforms, covering neuron and synapse implementations, learning rules, and mapping methods for spiking neural networks (SNNs). The authors contrasted biologically accurate models with reduced engineering-friendly models, emphasizing trade-offs required by scalable hardware. Their results stress the tradeoff between accuracy and efficiency in neuromorphic chip design.

Indiveri and Liu [2] surveyed memory and information processing strategies in neuromorphic circuits based on mixed-signal analog/digital circuits. They presented how event-driven dynamics and local memory emulate biological neurons and save energy. Their work showed that mixed-signal architectures successfully facilitate ultra-low power neuromorphic systems.

Davies et al. [3] evaluated the Intel Loihi platform, a many-core neuromorphic processor with on-chip learning capability. The paper benchmarked Loihi compared to legacy CPUs and GPUs, indicating dramatic improvements in latency and power efficiency. The authors concluded that neuromorphic cores are appropriate for real-time, adaptive edge tasks.

Neftci et al. [4] studied algorithm–hardware interface in neuromorphic computing, noting asynchronous and event-driven computation as vital for low-power intelligence. They stressed surrogate gradient techniques to train spiking networks effectively. Their research noted co-design between hardware and algorithms as being key to neuromorphic scalability.

Xiao [5] gave an extensive review of memristor materials and switching configurations pertinent to neuromorphic chips. The paper discussed device-level properties like endurance, variability, and analog programmability and related them to system-level issues. The author emphasized that trustworthy memristors are the key to next-generation large-scale neuromorphic systems.

Aguirre et al. [6] reviewed memristive artificial neural networks, focusing on crossbar array architectures and peripheral

circuits. In their work, they investigated variability issues and energy–accuracy trade-off in array-level computing. They concluded that optimized peripheral design and mapping strategies are crucial to enhance scalability.

Xu et al. [7] surveyed the application of memristors within ANN and SNN architectures. They emphasized device modeling, synaptic circuit integration, and dense array scaling as major on-chip learning enablers. The research emphasized the role of memristor non-volatility in designing efficient, compact neuromorphic processors.

Mishra et al. [8] investigated neuromorphic lifelong learning with an emphasis on how neuromorphic chips could enable lifelong learning without catastrophic forgetting. Their survey combined algorithmic designs with hardware capabilities and emphasized plasticity and local adaptation. The study showed the capability of neuromorphic systems towards autonomous, real-time adaptation.

Rathi et al. [9] offered a wide overview of advancements in spiking neuromorphic computing, highlighted CMOS and up-and-coming non-volatile memory technologies, and noted scalability and standardization to be the primary challenges. Their study emphasized the necessity of strong benchmarking tools to verify neuromorphic systems.

Wang et al. [10] analyzed memristor-based spiking neuromorphic systems, from device-level switching physics to neural dynamics. They surveyed candidate devices for SNN accelerators and analyzed their efficiency in emulating biological synapses. The research concluded that memristors provide special opportunities in developing brain-like AI hardware.

Xiao [11] discussed the evolution from memristor devices towards integrated neuromorphic chips. Their research considered reliability, hybrid CMOS–memristor integration, and design methodologies for deployment at the system level. The paper emphasized edge AI as the primary source of driving neuromorphic hardware development.

Min et al. [12] gave an overview of the ecosystem for memristor-based neuromorphic computing in terms of device, circuit, architecture, and algorithm layers. They promoted hybrid modeling approaches to mitigate the effects of device non-idealities. Their results reveal that cross-layer optimization is key to scalable neuromorphic systems.

Sokolov et al. [13] discussed oxide-based resistive RAM (RRAM) devices for neuromorphic systems. The paper compared various material systems and emphasized endurance, switching uniformity, and variability as the most important factors in synaptic emulation. According to the authors, material engineering is still a bottleneck for the reliability of neuromorphic hardware.

Al Abdul Wahid [14] also surveyed neuromorphic chip architectures, such as neuron models, synaptic topologies, and learning rules. Their survey aligned hardware designs with particular edge AI domains like robotics and IoT. The research also suggested evaluation metrics to unbiasedly compare neuromorphic implementations.

Muir et al. [15] surveyed commercialization routes for neuromorphic hardware, comparing research prototypes to industry-level systems. They concluded that fabrication complexity, market readiness, and scalability were the largest challenges. Their view highlighted the disparity between research in academia and product deployment in practice.

3. COMPARISONS OF VARIOUS WORKS

The literature surveyed reveals an unmistakable evolution from conceptual neuromorphic computing debates to tangible chip-level implementations focused on energy-efficient intelligence at the edge. Cornerstone papers like Indiveri and Liu (2015) and Neftci et al. (2018) set forth mixed-signal architectures and event-based models that laid the groundwork for brain-inspired computation, while Shrestha et al. (2022) and Rathhi et al. (2023) cited architectural reviews and benchmarking issues in spiking neural networks (SNNs). Device-level research such as Xiao (2023, 2024), Aguirre et al. (2024), and Xu et al. (2021) addressed crossbars and synapse circuits with memristors, emphasizing how in-memory

computing allows parallel operations with density but also revealing variability and endurance issues. Bringing devices to system dynamics, Wang et al. (2025) and Min et al. (2021) linked memristive switching to SNN efficiency, while suggesting hybrid modeling schemes to address non-idealities. On the application side, Davies et al. (2018) showed Intel's Loihi chip as a neuromorphic processor to provide low latency and high energy efficiency, and Mishra et al. (2023) built on this with ongoing learning frameworks to enable lifelong adaptability. Complementary work of Sokolov et al. (2021) and Al Abdul Wahid (2024) studied RRAM synapses and metrics for edge use cases, while Muir et al. (2025) focused on the gap between lab prototypes and deployable offerings for commercialization. Together, these papers point to neuromorphic chips as scalable, adaptive, and low-power von Neumann system alternatives, but identify issues to be resolved in fabrication reliability, variability management, hybrid CMOS–memristor integration, and system-level standardization toward eventual real-world uptake.

Table -1: Sample Table format

Ref. No.	Year	Authors (main)	Short Title / Topic	What the paper suggests / proposes
[1]	2022	Shrestha et al.	Neuromorphic computing survey	Reviews neuron/synapse models, SNN mapping, and trade-offs between biological realism and hardware.
[2]	2015	Indiveri & Liu	Memory & info processing in neuromorphic systems	Discusses mixed-signal circuits emulating brain-like neurons; emphasizes energy-efficient designs.
[3]	2018	Davies et al.	Intel Loihi chip	Describes Loihi neuromorphic processor with on-chip learning; shows energy & latency advantages.
[4]	2018	Neftci et al.	Event-driven neuromorphic intelligence	Proposes asynchronous/event-driven learning; stresses algorithm–hardware co-design.
[5]	2023	Xiao	Memristor materials & structures review	Reviews device-level materials, switching, and endurance issues for neuromorphic crossbars.
[6]	2024	Aguirre et al.	Memristive neural networks	Examines crossbar arrays, peripheral circuits, and energy–accuracy trade-offs.
[7]	2021	Xu et al.	Memristor-based neural networks	Reviews memristor device models, synapse circuits, and ANN/SNN integration for compact arrays.
[8]	2023	Mishra et al.	Continual learning in neuromorphic systems	Surveys lifelong learning, plasticity, and adaptation in neuromorphic chips.
[9]	2023	Rathhi et al.	SNN challenges & progress	Highlights scalability, benchmarking, and CMOS vs

Ref. No.	Year	Authors (main)	Short Title / Topic	What the paper suggests / proposes
				emerging memory for neuromorphic chips.
[10]	2025	Wang et al.	Memristor-based SNN systems	Links device physics to spiking dynamics; identifies device candidates for efficient SNNs.
[11]	2024	Xiao	Device-to-chip neuromorphic advances	Analyzes hybrid CMOS–memristor integration and reliability issues for edge AI systems.
[12]	2021	Min et al.	Memristor computing for neuromorphic platforms	Advocates cross-layer optimization to mitigate device non-idealities.
[13]	2021	Sokolov et al.	RRAM devices for neuromorphic applications	Reviews oxide-based RRAM, endurance, and variability challenges in synaptic emulation.
[14]	2024	Al Abdul Wahid	Survey of neuromorphic architectures	Maps neuron models, learning rules, and memory topologies to edge AI domains.
[15]	2025	Muir et al.	Commercialization challenges in neuromorphic HW	Identifies barriers from lab prototypes to products: fabrication complexity, scalability, markets.

4. DISCUSSIONS

Literature that is reviewed herein indicates that neuromorphic computing has evolved from theoretical models to a pragmatic chip realization for Edge-AI applications at a very fast pace. Pioneering works like Indiveri and Liu (2015) and Neftci et al. (2018) presented mixed-signal circuits and event-driven processing, building the foundations of brain-inspired processing. Expanding on these principles, Davies et al. (2018) showcased Intel's Loihi processor, which was shown to enable on-chip learning and real-time adaptation with significantly less power utilization than CPUs and GPUs. These contributions together set neuromorphic chips as an exciting candidate to break the von Neumann bottleneck.

The dominant research direction is the incorporation of memristor-based crossbar arrays to realize dense synaptic connectivity and low-power matrix–vector multiplications. Research such as Xiao (2023, 2024), Aguirre et al. (2024), and Xu et al. (2021) assures that memristors provide ultra-low power and small-scale designs that are suitable for large-scale neuromorphic hardware. Yet, the devices are also plagued by variability, endurance, and fabrication complexity issues, which detract from their reliability. Concurrently, Wang et al. (2025) and Min et al. (2021) show that integrating device physics with spiking neural networks (SNNs) can improve biological plausibility and efficiency, with hybrid modeling being suggested to counteract hardware flaws.

Applications of neuromorphic chips also exhibit their potential to promote real-time, adaptive intelligence in robotics, healthcare, and IoT. Mishra et al. (2023) build on this by emphasizing ongoing learning such that systems can learn without catastrophic forgetting, while Al Abdul Wahid (2024) and Sokolov et al. (2021) look at the need for benchmarking and material advancements to ensure long-term system

dependability. Lastly, Muir et al. (2025) bring attention to the disconnect between academic prototypes and commercialization, citing manufacturing constraints, CMOS–memristor integration hurdles, and standardization. Altogether, these works point toward neuromorphic chips as scalable, adaptive, and power-friendly replacements for conventional processors, while highlighting open challenges that direct future research in hybrid integration.

5. CONCLUSION

This review has discussed the evolution of neuromorphic computing from its theoretical roots to state-of-the-art chip-level designs that have the potential to redefine intelligent information processing at the edge. In the studies reviewed, the common thread is: neuromorphic chips, especially those incorporating memristor crossbar arrays, bring a special combination of in-memory computation, parallelism, and ultra-low power consumption that tackles head-on the latency, energy, and scalability issues of von Neumann architectures. Such platforms have been shown by researchers to support spiking neural networks, adaptive learning, and real-time inference, and they can be used for applications such as robotics, healthcare, IoT, and smart infrastructure where responsiveness and efficiency are essential. Hybrid solutions with the combination of device innovations with algorithm–hardware co-design further enhance adaptability, and continual learning frameworks push neuromorphic systems towards lifelong intelligence.

At the same time, the literature highlights key challenges that must be addressed before widespread deployment. Device variability, fabrication complexity, and limited endurance constrain memristor reliability, while system-level bottlenecks arise in scaling to larger networks and integrating heterogeneous circuits. Commercialization barriers, including hybrid CMOS–memristor integration and standardization of evaluation benchmarks, remain open areas for research. Nevertheless, the

collective evidence reviewed here shows that neuromorphic chips are moving steadily from conceptual prototypes toward practical tools for energy-efficient, real-time, and adaptive edge computing. Continued advances in device engineering, 3D crossbar architectures, and biologically inspired learning mechanisms are likely to deliver the next generation of brain-inspired hardware platforms, establishing neuromorphic chips as a cornerstone of future intelligent systems.

REFERENCES

1. A. Shrestha and A. Mahmood, "Review of Neuromorphic Computing Models and Hardware Platforms," *Journal of Systems Architecture*, vol. 122, pp. 102–118, 2022.
2. G. Indiveri and S.-C. Liu, "Memory and Information Processing in Neuromorphic Systems," *Proceedings of the IEEE*, vol. 103, no. 8, pp. 1379–1397, 2015.
3. M. Davies, N. Srinivasa, T.-H. Lin, et al., "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018.
4. E. O. Neftci, H. Mostafa, and F. Zenke, "Surrogate Gradient Learning in Spiking Neural Networks," *Nature Machine Intelligence*, vol. 1, no. 1, pp. 1–12, 2018.
5. T. Xiao, "Memristor Materials and Device Structures for Neuromorphic Computing: A Review," *Advanced Electronic Materials*, vol. 9, no. 4, 2023.
6. F. Aguirre, P. K. Panda, S. B. Eryilmaz, et al., "Memristive Neural Networks: From Device Physics to Large-Scale Architectures," *Nature Communications*, vol. 15, pp. 1–15, 2024.
7. W. Xu, Y. Wang, and J. Yang, "Advances in Memristor-Based Neural Networks: Devices, Circuits, and Applications," *Frontiers in Nanotechnology*, vol. 3, pp. 45–59, 2021.
8. S. Mishra, A. Kumar, and R. Tiwari, "Neuromorphic Continual Learning: Algorithms, Architectures, and Nano-Device Solutions," *Frontiers in Neuroscience*, vol. 17, pp. 1–12, 2023.
9. N. Rath, R. Thakur, and A. Basu, "Challenges and Progress in Neuromorphic Computing Based on Spiking Neural Networks," *IEEE Access*, vol. 11, pp. 84521–84535, 2023.
10. H. Wang, X. Li, and Y. Zhang, "Memristor-Based Spiking Neuromorphic Systems: A Review," *Nanomaterials (MDPI)*, vol. 15, no. 3, pp. 331–348, 2025.
11. L. Xiao, "From Memristive Devices to Neuromorphic Chips: Progress and Future Directions," *Journal of Applied Physics*, vol. 136, no. 2, pp. 021101–021120, 2024.
12. K. Min, J. Lee, and S. Park, "Memristor Computing for Neuromorphic Systems: Cross-Layer Co-Design Approaches," *Applied Sciences*, vol. 11, no. 19, pp. 9055–9067, 2021.
13. A. S. Sokolov, P. Karpov, and M. Kozodaev, "Materials and RRAM Devices for Non-Volatile Memory and Neuromorphic Applications," *Nanotechnology*, vol. 32, no. 24, pp. 243001–243015, 2021.
14. A. A. Wahid, "Survey of Neuromorphic Architectures and Edge AI Applications," *ACM Computing Surveys*, vol. 56, no. 3, pp. 1–27, 2024.
15. D. R. Muir, F. Zenke, and T. Pfeil, "Commercialization Challenges and Opportunities for Neuromorphic Hardware," *Nature Communications*, vol. 16, pp. 1–10, 2025.