Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 1s (2025)

# Neuromorphic Computing Chips for Low-Power Edge AI Applications

**Dr Jyoti G**

Associate Professor in Electronics, Government Science College, Nrupathunga University, N. T. Road, Bangalore-560001, Karnataka, India.

***Article History:***

***Received: 02-10-2024***

***Revised: 25-11-2024***

***Accepted: 20-01-2025***

**Abstract**

This paper highlights why neuromorphic computing chip is promising to meet low-power edge AI application. The primary objective is to examine their effectiveness, real-time processing potential, comparative efficiency to standard accelerators and how viable they may be in an industrial setting. The data used was a secondary data method and made use of peer-reviewed journal articles, conference papers and industry reports to compile validated statements on architecture, power metrics and the application outcomes. The results also demonstrate that neuromorphic processors based on spiking neural networks technology, in-memory processing, and using memristor crossbar arrays have much higher energy efficiency and an order of magnitude better response times (milliseconds) than GPU and CNN accelerators. In addition, they allow on-chip learning and continuous adaption with power budgets below 10 mW, thus are ideally suited to IoT, robotic, healthcare wearables, and industrial IoT. The study in general concluded that, the neuromorphic chips are a significant tool to support the overall realisation of sustainable, adaptive and real-time edge intelligence in low-power systems.

**Keywords-** Neuromorphic, Chips, Edge AI, Power, Energy efficiency, Accelerators, Processing, Spiking Neural Networks (SNNs), IoT, Memristor

## Introduction

Neuromorphic computing chips are equally becoming revolutionary when it comes to low-power edge AI functions. Motivated by human brains, these chips behave similar to human brains in that they process information by using event-driven and parallel frameworks. In contrast to Von Neumann architecture, neuromorphic hardware has memory and computational cointegration, thereby saving a lot of energy, as well as latency. This qualifies them well to carry out real-time applications like image recognition, speech processing, anomaly detection at edge where devices are normally power and bandwidth challenged. Firms such as Intel, IBM, and BrainChip, are developing neuromorphic CPUs that can operate on power at the milliwatt level. The need to find a solution to efficient, scalable deployment of AI into various industries increases, and neuromorphic chips will provide an attractive way to generate, as well as ensure energy efficiency and sustainability.

**Objectives**

- To evaluate the efficiency of neuromorphic chips in reducing power consumption for edge AI workloads.
- To analyze the performance of neuromorphic architectures in real-time data processing tasks.
- To compare neuromorphic hardware with conventional AI accelerators in terms of scalability and latency.
- To explore industrial applications of neuromorphic chips for sustainable edge intelligence

**Literature Review**

Vaibhav (2025):

Vaibhav has a brainy idea of neuromorphic-style VLSI that would suit small sensor nodes with an ultra-low power envelope. According to the design, it combines in-memory processing, and event-driven spiking cores, reducing the data movement and idle power. A wake-on-event front end suppresses sensor noise and initiates inferencing within milliseconds to produce anomaly detection and keyword spotting results with microsecond response times. Compact STDP implementation using quantized synapses enables on-chip learning even in the absence of access to the cloud. In contrast to small CNN accelerators, the chip is claimed to achieve a better energy efficiency and duty-cycled lifetimes, which are important to always-on edge devices and discontinuous energy harvesting.

Das (2024):

A systematic literature review of neuromorphic computing in edge AI is presented by Das in terms of its hardware, algorithms and application domains. The taxonomy cuts across Loihi, TrueNorth, and Akida platforms; applications of the taxonomy include keyword search, human activity, and predictive maintenance. Reported measures further focuss on energy per inference, and TOPS/W, where event-driven SNNs continually outperform dense DNN baselines, under power constraints. There were found gaps in the immaturity of toolchains, the lack of available SNN datasets, and robustness under non-stationary edge conditions. The review does support sensor-algorithm-hardware co-design, learning within the chip to personalize a system and mixed-signal designs to unleash ultra-low-power and real-time autonomy.

Ye et al. (2023):

Ye and colleagues summarize advances in the design of low-power AIoT chips with a focus on near-sensor computing, approximate arithmetic, and non-volatile memory crossbars (RRAM/PCM) to compute in-situ MAC operations. They detail the system-level co-optimization: quantization-aware training, sparsity, and model pruning, providing energy comprehension purposes by an order of amount, as compared with basic accelerators. The

survey describes clock-/power-gating, dynamic voltage/frequency scaling and event-driven dataflows which match well with spiking networks. An example of a practical AIoT SoC going mainstream will use 10-100x efficiency factors and sub-millisecond latency in interpreting vision and audio signals. There are outstanding issues with device variability, wear of NVM arrays and standardised metrics in edge cases.

Nwakanma, Abosede, et al. (2021):

NeuroEdge is an asynchronous neuromorphic computing concept in edge AI proposed by NeuroEdge co-authors that combines asynchronous event routing with a crossbar synaptic array and lightweight digital control. The paper justifies the use of neuromorphic in power-constrained wearable, UAVs and smart cameras with streaming sensory spikes being able to provide sparse and low latent inference. They speak about programmer-friendly toolchain and neuromorphic middleware as well as the trade-off in architecture precise versus energy, local learning versus stability. The migration is framed into a workflow that goes from DNN-centered to application-specific to energy efficient to always-on with a focus on event driven SNNs.

## Methodology

This study focuses on addressing secondary data, which includes using peer-reviewed journals, conference proceedings, and industry reports on the topics of neuromorphic computing and edge AI. Secondary data allows access to large volumes of empirical research results, performance indicators and cross-industry experiences without experimental prototypes which require significant cost and can consume much time. The works by Lin et al. (2021), Luo et al. (2023), and Xia et al. (2025) present validated information on the power efficiency, latency, and architectural design on which it becomes possible to compare platforms. The advantage of this process is that there is a synthesis of various views, there are technological gap issues involved, and there is reliability since different established outcomes were triangulated. It also reduces bias on the part of the researcher by using published data only. Generally, secondary data make use more accurate, feasible, and extensive, which is appropriate to the assessment of low-power edge AI in the use of the neuromorphic chip.

## Result and Discussion

### *Energy Efficiency of Neuromorphic Chips in Edge AI Workloads*

Neuromorphic chips realized enhanced energy efficiency relative to conventional accelerators, in particular due to event-driven architecture and in-memory computing. The implementation of neuromorphic systems is shown by Luo et al. (2023) to achieve 90 percent less energy consumption in deep learning than GPUs and tens of TOPS/W in inference applications. Liu et al. (2021) also indicate how neuromorphic engineering takes advantage of the spiking neural networks (SNNs) that use sparse spikes rather than dense activations, which reduces dynamic power. In line with this, Putra et al. (2025) demonstrate that physics-based on-chip learning on

commodity neuromorphic processors circumvents longer and energy-consuming cloud retraining to enable autonomic adaptation even under <10 mW power budget.

| Processor Type | Power Consumption (W) | Efficiency (TOPS/W) | Energy Reduction (%) | Reference |
|---|---|---|---|---|
| GPU (NVIDIA Jetson Nano) | 10 W | 5 TOPS/W | Baseline | Lin et al., 2021 |
| Neuromorphic Chip (Loihi 2) | 0.02 W | 60 TOPS/W | 90% | Luo et al., 2023 |
| Memristor-based Accelerator | 0.01 W | 75 TOPS/W | 95% | Xia et al., 2025 |
| Neuromorphic with On-chip STDP | 0.005 W | 80 TOPS/W | 96% | Putra et al., 2025 |

**Table 1: Energy Efficiency Comparison of Neuromorphic Chips and Conventional Accelerators**

In neuromorphic hardware implementation (see Xia et al., 2025) near-memory computing at von Neumann bottleneck-free femtojoule levels is achieved by using a memristor medium. In the same way, Jebali et al. (2024) introduce solar cell powered memristor BNN accelerators operating at a low power consumption level, which is ideal to support inference in remote sensors. These developments identify that neuromorphic chips, in contrast to CNN accelerators (Lin et al., 2021), balance latency and energy consumption, which is reasonable to be supported by battery-based systems, including IoT nodes, drones, and wearables, where the number of watts defines the feasibility of the system.

***Real-Time Data Processing Capabilities of Neuromorphic Architectures***

Artificial intelligence Neuromorphic architectures are more than capable of real-time processing because they emulate the parallel, asynchronous dynamics of biological neurons. The results of Lin et al. (2021) show that a reduction of more than 50 percent in inference latency of ultra-small edge accelerators with neuromorphic integration compared to conventional CNN accelerators achieves real-time vision recognition. Ivkovic and (2023) demonstrate that cognitive SNN computing can be done within MCUs with integrated NPU/GPU accelerators in disconnected environments at millisecond decision time scales. Putra et al. (2025) show on-chip learning driven spikes which support dynamic anomaly detection and avoid retraining delays to enable predictive maintenance in industrial sensors. Jussinmaki et al. (2025) note STM and nRF microcontrollers with neuromorphic-inspired optimizations that can allow these systems to act within a range of 1-5ms (a possible 1-5ms response in robotics and autonomous navigation).

| Platform/Architecture | Task Type | Latency (ms) | Response Speed Gain (%) | Reference |
|---|---|---|---|---|
| Conventional CNN Accelerator | Image Recognition | 12 ms | Baseline | Lin et al., 2021 |
| Neuromorphic MCU (Cognitive SNN) | Sensor Event Detection | 5 ms | 58% | Ivković & Ivković, 2023 |
| Neuromorphic Microcontroller | Robotics Navigation | 3 ms | 75% | Jussinmäki et al., 2025 |
| On-chip SNN Learning | Anomaly Detection | 2.5 ms | 79% | Putra et al., 2025 |

**Table 2: Real-Time Processing Performance of Neuromorphic Architectures**

Neuromorphic systems can accomplish this by implementing asynchronous spike routing, personalised memory access, and on-board inference, minimising dependence on cloud inference. According to Sigala (2025), these features are essential when an application requires low latencies or has high safety-related demands such as autonomous drones or wearable health monitors (where latency values above 10 ms may cause safety problems). Together, these results verify neuromorphic architectures as being the best enablers of real-time intelligence in delimited edge cases, beating traditional accelerators in the domains of both latency and task flexibility.

***Comparative Performance of Neuromorphic and Conventional AI Accelerators***

Comparative analysis shows that compared to traditional AI accelerators, the neuromorphic hardware works much better in energy-limited settings. According to Loihi neuromorphic processors, utilizing such devices 15x LE than GPUs on sparse workloads, with only a slight difference in accuracy when used in speech recognition. Lin et al. (2021) provide a comparison of low-power CNN accelerators and neuromorphic chips, and state that the latter may exceed CNN accelerator performance in 50-100 TOPS/W as MAC operations can be done on crossbar memristors that are integrated in the chips. According to Putra et al. (2025), since on-chip STDP learning reduces the memory transfer costs, it contrasts with GPU which show high memory access overheads.

| Accelerator Type | Power (W) | Accuracy (%) | Energy Efficiency (TOPS/W) | Improvement Factor | Reference |
|---|---|---|---|---|---|
| GPU (Edge Inference) | 15 W | 92% | 10 TOPS/W | Baseline | Luo et al., 2023 |

| | | | | | |
|---|---|---|---|---|---|
| CNN Edge Accelerator | 2 W | 91% | 50 TOPS/W | 5× | Lin et al., 2021 |
| Neuromorphic (Loihi) | 0.1 W | 90% | 150 TOPS/W | 15× | Luo et al., 2023 |
| Memristor BNN Chip | 0.05 W | 89% | 180 TOPS/W | 18× | Jebali et al., 2024 |

**Table 3: Comparative Performance of Neuromorphic vs Conventional Accelerators**

Jebali et al. (2024) further support this through the use of binarized neural networks based on memristors that provide sustained inference with milliwatts of power and GPU performance requires tens of watts. According to Lajlan and Ibrahim (2022), TinyML systems to run CNN inferences are not very adaptable and do not measure up to the SNN neuromorphic systems in terms of self-learning capacity on the continuous sensory inputs. All in all, neuromorphic processors offer an improvement by an order of magnitude in energy-delay product and improved scalability of resource-constrained embedded devices, and by extension as a disruptive technology to conventional accelerators in low-power autonomous AI use cases.

***Industrial Applications of Neuromorphic Chips for Sustainable Edge Intelligence***

Neuromorphic chips are in motion across sectors that need edge brilliance that is both enduring and ever-present. The use of their counterparts in smart cities is featured in a paper by Galia (2025): energy-efficient anomaly detection systems target to attain a sensor lifetime extension range of 5-10 years under solar power consumption. TinyML and neuromorphic fusion-based technologies allow non-stop tracking of ECG to detect arrhythmias at less than 5mW (Lajlan and Ibrahim, 2022). In industrial IoT, Putra and et al. (2025) cite neuromorphic processors as deployed to predictive maintenance, detecting motor anomalies in less than 3 ms, avoiding the costly downtime.

| **Industry Sector** | **Application Example** | **Power Requirement (mW)** | **Performance Metric** | **Reference** |
|---|---|---|---|---|
| Healthcare | ECG Monitoring Wearable | 5 mW | Continuous arrhythmia detection | Alajlan & Ibrahim, 2022 |
| Industrial IoT | Motor Fault Detection | 8 mW | Anomaly detected in <3 ms | Putra et al., 2025 |
| Robotics | Autonomous Navigation MCU | 12 mW | 40% energy savings | Jussinmäki et al., 2025 |

| | | | | |
|---|---|---|---|---|
| Smart Cities | Anomaly Detection Sensors | 6 mW | 10-year operational lifetime | Sigala, 2025 |

**Table 4: Industrial Applications of Neuromorphic Chips**

Uses include autonomous robotics, where, in an alternative embodiment, neuromorphic microcontrollers further save energy by 40 per cent over systems powered by GPUs (Jussinmaki et al., 2025). UAV and drone use are also promoted, as sparse spiking minimises the onboard computational cost, and UAV flight endurance is increased by 20 30%. With reference to materials science, Xia et al. (2025) prioritise the in-size energy harvesting and embedded intelligence of smart grids using memristor-based neuromorphic devices. Taken together, these results validate that neuromorphic chips are not simply new technology demonstrators but viable solutions to the intelligent and sustainable end-to-end device and infrastructure, such as those seen in healthcare, robotics, industrial IoT, and urban infrastructure, where power efficiency, real-time responsiveness, and system life are paramount to deployment.

## Conclusion

In this study, the researchers have found that neuromorphic computing chips are a disruptive innovation to further provide significant energy savings, latency improvements, as well as real-time responsiveness to edge AI applications. Secondary sources provide evidence that programmable accelerators based on event-driven SNNs, in-memory architectures, and memristor-based designs show 10-15x energy gains over conventional accelerators, which enable long-term sustainable intelligence on the IoT, robotics, health monitoring, and other industrial systems. In contrast to conventional GPU and CNN-based systems, neuromorphic systems allow continuous learning and inference at the milliwatt power consumption level, a quality suitable to battery-powered scenarios, or disconnected settings. The paper also has very good industrial applicability in smart cities, UAVs and wearable devices. Thus, neuromorphic chips are one of the most important technologies to implement scalable, adaptive and sustainable AI at the edge of the network.

## References

1. Alajlan, N.N. and Ibrahim, D.M., 2022. TinyML: Enabling of inference deep learning models on ultra-low-power IoT edge devices for AI applications. *Micromachines*, *13*(6), p.851.
2. Das, R.S., 2024. Emerging Neuromorphic Computing for Edge AI Application: A Systematic Literature Review. *Journal of Technological Innovations*, *5*(1).
3. Ivković, J. and Ivković, J.L., 2023. Exploring the potential of new AI-enabled MCU/SOC systems with integrated NPU/GPU accelerators for disconnected Edge computing applications: towards cognitive SNN Neuromorphic computing. In *LINK IT> EdTech International Scientific Conference. Belgrade* (pp. 12-22).

4. Jebali, F., Majumdar, A., Turck, C., Harabi, K.E., Faye, M.C., Muhr, E., Walder, J.P., Bilousov, O., Michaud, A., Vianello, E. and Hirtzlin, T., 2024. Powering AI at the edge: A robust, memristor-based binarized neural network with near-memory computing and miniaturized solar cell. *Nature Communications*, *15*(1), p.741.
5. Jussinmäki, L., Plosila, J. and Haghbayan, H., 2025. Bringing Edge AI to low-power nRF and STM microcontrollers. *Robotics and Autonomous Systems*.
6. Lin, W., Adetomi, A. and Arslan, T., 2021. Low-power ultra-small edge AI accelerators for image recognition with convolution neural networks: Analysis and future directions. *Electronics*, *10*(17), p.2048.
7. Liu, D., Yu, H. and Chai, Y., 2021. Low-power computing with neuromorphic engineering. *Advanced Intelligent Systems*, *3*(2), p.2000150.
8. Luo, T., Wong, W.F., Goh, R.S.M., Do, A.T., Chen, Z., Li, H., Jiang, W. and Yau, W., 2023. Achieving green ai with energy-efficient deep learning using neuromorphic computing. *Communications of the ACM*, *66*(7), pp.52-57.
9. Nwakanma, C.I., Kim, J.W., Lee, J.M. and Kim, D.S., 2021. Edge AI prospect using the NeuroEdge computing system: Introducing a novel neuromorphic technology. *ICT Express*, *7*(2), pp.152-157.
10. Putra, R.V.W., Wickramasinghe, P. and Shafique, M., 2025. Enabling Efficient Processing of Spiking Neural Networks with On-Chip Learning on Commodity Neuromorphic Processors for Edge AI Systems. *arXiv preprint arXiv:2504.00957*.
11. Sigala, N.S., 2025. Edge AI for Energy-Efficient Computing: A Systematic Review of Strategies, Challenges, and Future Directions. *International Journal of Business & Computational Science*, *2*(1).
12. Vaibhav, V.G., 2025. A Neuromorphic-Inspired, Low-Power VLSI Architecture for Edge AI in IoT Sensor Nodes. *Journal of Microelectronics and Solid State Devices*, *12*(2), pp.41-47p.
13. Xia, Z., Sun, X., Wang, Z., Meng, J., Jin, B. and Wang, T., 2025. Low-power memristor for neuromorphic computing: From materials to applications. *Nano-Micro Letters*, *17*(1), p.217.
14. Ye, L., Wang, Z., Jia, T., Ma, Y., Shen, L., Zhang, Y., Li, H., Chen, P., Wu, M., Liu, Y. and Jing, Y., 2023. Research progress on low-power artificial intelligence of things (AIoT) chip design. *Science China Information Sciences*, *66*(10), p.200407.
15. Ye, L., Wang, Z., Jia, T., Ma, Y., Shen, L., Zhang, Y., Li, H., Chen, P., Wu, M., Liu, Y. and Jing, Y., 2023. Research progress on low-power artificial intelligence of things (AIoT) chip design. *Science China Information Sciences*, *66*(10), p.200407.