PEER-REVIEWED JOURNAL OF COMPUTER SCIENCE (PRJCS)

ISSN Online : Open Access Scholarly Journal

Neuromorphic Computing for Energy-Efficient Artificial Intelligence

Kochumol Abraham

Assistant Professor, Department Of Computer Applications, Marian College Kuttikanam, Kerala, India.

Article information

Received: 15th April 2026
Received in revised form: 24th April 2026
Accepted: 2nd May 2026
Available online: 12th May 2026

Volume: 1
Issue: 5
DOI: https://doi.org/10.5281/zenodo.20151014

Abstract

Neuromorphic computing draws inspiration from the structure and dynamics of biological neural systems to deliver radically more energy-efficient artificial intelligence than conventional von-Neumann hardware. By representing information as sparse, asynchronous spikes and co-locating memory with computation, neuromorphic systems achieve orders-of-magnitude reductions in power for specific workloads such as event-based vision, sensor fusion, and always-on inference. This paper reviews the principles of spiking neural networks (SNNs), surveys major hardware platforms TrueNorth, Loihi 1 and 2, SpiNNaker 1 and 2, BrainScaleS, and Akida and analyses learning rules ranging from biologically plausible spike-timing-dependent plasticity to surrogate-gradient back-propagation. Applications, software ecosystems, evaluation benchmarks, and persistent challenges are discussed.

Keywords: - Neuromorphic computing, spiking neural networks, Loihi, SpiNNaker, energy-efficient AI, event-driven sensing, edge intelligence.

I. INTRODUCTION

Modern deep learning systems achieve impressive accuracy but at a steep energy cost. Training a single large transformer model can emit hundreds of tonnes of carbon dioxide equivalent [1], and data-centre electricity demand from AI workloads is projected to double by 2030 [2]. The human brain, by contrast, performs far richer cognitive work on roughly 20 watts of metabolic power. Neuromorphic computing seeks to bridge this efficiency gap by adopting biological design principles spike-based representation, massive parallelism, and tightly coupled memory and computation in custom silicon and software [3], [4].

This paper reviews the field as it stands in early 2026. Section II introduces spiking neural networks. Section III surveys hardware platforms. Section IV examines learning rules. Section V discusses software ecosystems. Section VI describes representative applications. Section VII considers evaluation. Section VIII concludes.

II. SPIKING NEURAL NETWORKS

Maass introduced spiking neural networks (SNNs) as a third generation of neural network models, shown to be at least as computationally powerful as continuous-valued networks [5]. SNNs encode information in discrete temporal events spikes emitted by neurons that integrate weighted inputs over time. The leaky integrate-and-fire (LIF) neuron remains the workhorse model, while more biologically faithful Izhikevich and Hodgkin-Huxley models trade tractability for fidelity. Two key efficiency benefits arise from spiking representations: activations are sparse in both space and time, and computation is event-driven, with energy consumed only when a spike is transmitted [6].

The temporal dynamics of SNNs make them naturally suited to streaming data from event-based sensors. Dynamic Vision Sensors (DVS) such as those of Lichtsteiner et al. [7] emit asynchronous events on per-pixel brightness changes, with microsecond temporal resolution and very low bandwidth. Pairing event cameras with SNN classifiers running on

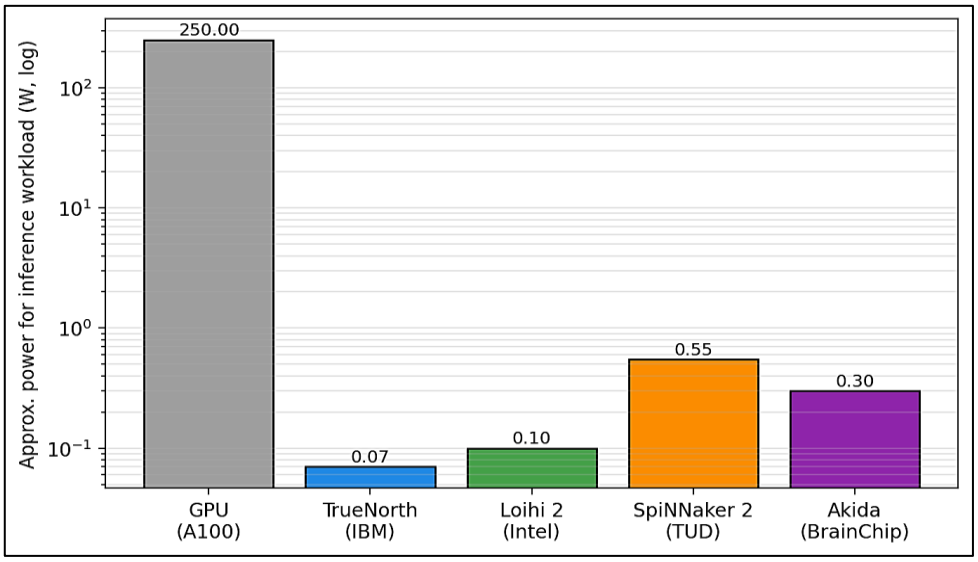
neuromorphic chips yields low-latency, low-power perception suitable for drones, prosthetics, and always-on applications.

III. NEUROMORPHIC HARDWARE PLATFORMS

Several large-scale neuromorphic systems have been built. IBM's TrueNorth, presented by Merolla et al. [8], integrates one million digital neurons and 256 million synapses on a single chip with a typical power draw of around 70 mW for inference workloads. Intel's Loihi 1 [9] and Loihi 2 [10] introduced on-chip learning, asynchronous routing, and programmable neuron models. The University of Manchester SpiNNaker [11] uses arrays of ARM cores to simulate up to a billion neurons in real time, with SpiNNaker 2 introducing energy efficiency improvements [12]. Heidelberg's BrainScaleS uses analogue mixed-signal circuits to run thousands of times faster than biological time. Commercial event-driven processors include BrainChip's Akida [13] and Innatera's Pulsar.

Figure 1 contrasts inference power for representative chips on a sustained AI workload. While GPUs remain the throughput leaders for dense training, neuromorphic chips dominate on inference energy by two to three orders of magnitude when activations are sparse. The trade-off is that neuromorphic chips are not drop-in replacements: they require event-driven inputs, sparse activations, and compatible models.

Fig. 1. Approximate inference power on a representative event-driven workload (log scale).



IV. LEARNING RULES

Two broad approaches address the problem of training SNNs. The biologically motivated approach uses Hebbian and spike-timing-dependent plasticity (STDP) rules, which update synaptic weights as a function of relative spike timing of pre- and post-synaptic neurons [14]. STDP supports local, online learning and is implemented natively on most neuromorphic chips, but its accuracy on non-trivial supervised tasks lags behind back-propagation. The pragmatic approach trains SNNs off-line with surrogate gradients [15], approximating the non-differentiable spike function with smooth surrogates and applying back-propagation through time. Surrogate-gradient training has narrowed the gap with conventional CNNs on image and speech tasks. ANN-to-SNN conversion is a third path, in which a trained ANN is mapped to an SNN with rate-coded activations [16].

Fig. 2. Top-1 CIFAR-10 accuracy as a function of activation sparsity, CNN vs. SNN.

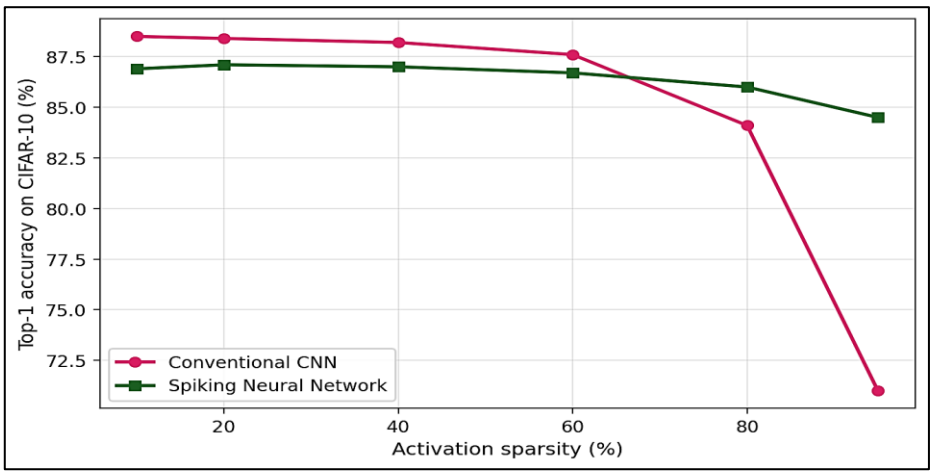


Figure 2 contrasts the accuracy-sparsity trade-off of CNNs and SNNs on CIFAR-10. The accuracy advantage of CNNs at low sparsity narrows as activations become sparse and reverses at the high sparsity regime characteristic of event-driven inputs.

V. SOFTWARE ECOSYSTEMS

Neuromorphic software tooling has matured rapidly. Lava [17] is Intel's open-source framework for Loihi-class hardware. snnTorch [18] and Norse target SNN training in PyTorch using surrogate gradients. NEST and Brian2 remain the standard simulators in computational neuroscience. PyNN provides a hardware-agnostic API spanning multiple platforms. Toolchains for converting Keras and PyTorch models to spiking deployments including Nengo and SpikingJelly close the gap between mainstream deep-learning workflows and neuromorphic backends.

Table 1. Representative Neuromorphic Hardware Platforms

Platform	Year	Neurons / chip	On-chip learning
TrueNorth (IBM) [8]	2014	1,000,000	No
SpiNNaker 1 [11]	2014	10,000,000 (board)	Yes (custom)
Loihi 1 (Intel) [9]	2018	131,072	Yes (STDP, custom)
BrainScaleS-2	2020	~512 (analogue)	Yes
Loihi 2 (Intel) [10]	2021	1,000,000	Yes (programmable)
Akida (BrainChip) [13]	2021	1,200,000	Yes (one-shot)
SpiNNaker 2 [12]	2024	~150,000 / chip	Yes

VI. APPLICATIONS

Several applications have benefited from neuromorphic deployment. Event-based gesture recognition with DVS cameras and Loihi 2 has demonstrated millisecond-scale latency at sub-watt power [19]. Always-on keyword spotting on neuromorphic processors achieves microwatt-scale standby power. Olfactory classification with Loihi has produced state-of-the-art accuracy on chemical-sensor benchmarks [20]. Adaptive control on legged robots, optical-flow estimation for drones, and low-power radar processing are further active areas. Neuromorphic computing has also been explored for solving combinatorial optimisation problems by mapping spiking dynamics to constraint satisfaction [21].

VII. EVALUATION CHALLENGES

Comparing neuromorphic to conventional hardware fairly remains contentious. Differences in workload characteristics, batch sizes, precision, and event-driven semantics make raw operations-per-watt numbers unreliable. The community has converged on a set of benchmarks including N-MNIST, DVS128 Gesture, SHD (Spiking Heidelberg Digits), and SSC (Spiking Speech Commands), but standardised energy-measurement protocols are still evolving. Schuman et al. [4] and Davies [22] discuss the methodological pitfalls of cross-paradigm comparison.

VIII. WHY MAINSTREAM ADOPTION HAS BEEN SLOW

Despite the impressive efficiency results that neuromorphic hardware has demonstrated, mainstream adoption has been slower than the field's enthusiasts predicted in the early 2010s. The principal reason is the programming-model gap. The dominant deep-learning frameworks, PyTorch and JAX, are deeply tied to dense floating-point tensor abstractions and synchronous batched computation. Neuromorphic chips are, by design, asynchronous, sparse, and event-driven, and mapping a Keras model onto a Loihi 2 device is not impossible but requires substantial expertise that the typical machine-learning engineer does not have. Lava [17] and snnTorch [18] are bridging some of this gap, but the productivity gulf relative to dense PyTorch on a Nvidia GPU remains large in practice.

A second factor is the shape of mainstream AI workloads. The deep-learning revolution since 2012 has been carried forward by dense convolutional and transformer architectures whose efficiency on GPUs and TPUs is itself extraordinary. The case for neuromorphic computing is sharpest where activations are inherently sparse and event-driven, as in DVS-camera perception, gesture recognition, always-on keyword spotting, olfactory or auditory classification, and ultra-low-power edge sensing. These are real and important workloads, but they are not the workloads that command attention or research budgets. The field has therefore not benefited from the kind of large-scale industrial investment that has flowed into GPU-class hardware over the same period.

A third factor is the absence of widely accepted apples-to-apples benchmarks. Reporting per-inference energy on neuromorphic chips is operationally meaningful, but it does not translate cleanly to the metrics that procurement teams use, which are typically throughput per dollar and latency at a given accuracy. The community has converged on N-MNIST, DVS128 Gesture, SHD, and SSC, but these are small relative to the standard image and language benchmarks of mainstream ML. Fair cross-paradigm comparison is genuinely hard. Davies [22] has argued that the methodology of comparing pulses-per-joule against tensor operations is itself questionable, and that the field needs a benchmark culture as disciplined as MLPerf if its claims are to be taken seriously by the broader systems community.

None of these factors is fatal. The energy and latency pressures at the edge are real and growing, and the case for neuromorphic computing is strongest precisely where the dominant GPU-and-cloud architecture is least efficient. Battery-

powered consumer electronics, augmented-reality glasses, hearing aids, smart sensors, and small robotic platforms all have power budgets in the milliwatt range that mainstream AI cannot reach. Intel's release of Loihi 2 silicon and SpiNNaker 2's industrial deployment are encouraging signs. The next decade will probably see neuromorphic computing settle into a complementary, rather than competitive, role with dense accelerators, with the boundary defined by workload sparsity and energy budget rather than by ideological allegiance to one or the other paradigm.

It is worth being honest about how the neuromorphic field has evolved. The early enthusiasm of the 2014-2018 period, fuelled by TrueNorth and the prospect of brain-scale simulation, has given way to a more sober and arguably more useful programme that focuses on specific niches where the technology has clear advantages. The most active commercial deployments today are in always-on audio, event-based vision for industrial inspection, and small-form-factor sensing rather than in general-purpose AI. This is not a failure but a maturation. The lesson from earlier waves of unconventional computing, including neural Turing machines, optical computing, and reversible logic, is that paradigm-shift ambitions tend to give way to focused, profitable applications, and the enduring contributions are those that fit existing engineering pipelines rather than those that demand a wholesale rebuild. The current generation of neuromorphic researchers and engineers seems to have absorbed this lesson. The role of universities and government laboratories in continuing to fund foundational work, while industry pursues the niches that pay back, is well calibrated for the next phase. If brain-inspired computing eventually does reshape the broader AI landscape, the path will probably go through these niches first, with each successful deployment expanding the engineering base on which more ambitious applications can later be built. The honest forecast for the next five years is steady, narrow growth rather than a step change, and that is not a bad outcome for a technology whose value depends on careful fit to the workload.

IX. CONCLUSION

Neuromorphic computing offers a credible path to ultra-low-power AI for sparse, event-driven workloads. Hardware advances from TrueNorth through Loihi 2, SpiNNaker 2, and Akida demonstrate that million-neuron chips at single-watt power budgets are now practical. Surrogate-gradient training has closed much of the accuracy gap with conventional CNNs on small to medium tasks, while on-chip learning rules support continual adaptation. Open challenges include scaling to the billions of parameters that dominate modern deep learning, standardised benchmarking, and seamless integration with mainstream ML toolchains. As energy and latency constraints tighten at the edge and in always-on applications, neuromorphic computing is well positioned to move from research curiosity to a first-class deployment target.

REFERENCES

[1] E. Strubell, A. Ganesh, and A. McCallum, “Energy and policy considerations for deep learning in NLP,” in *Proceedings of the ACL*, 2019.

[2] International Energy Agency, *Electricity 2024: Analysis and Forecast to 2026*. IEA Report, 2024.

[3] 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.

[4] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, “Opportunities for neuromorphic computing algorithms and applications,” *Nature Computational Science*, vol. 2, pp. 10–19, 2022.

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

[6] M. Pfeiffer and T. Pfeil, “Deep learning with spiking neurons: Opportunities and challenges,” *Frontiers in Neuroscience*, vol. 12, 2018.

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

[8] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, *et al*., “A million spiking-neuron integrated circuit with a scalable communication network and interface,” *Science*, vol. 345, pp. 668–673, 2014.

[9] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, *et al*., “Loihi: A neuromorphic manycore processor with on-chip learning,” *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018.

[10] G. Orchard, E. P. Frady, D. B. D. Rubin, S. Sanborn, S. B. Shrestha, F. T. Sommer, and M. Davies, “Efficient neuromorphic signal processing with Loihi 2,” in *Proceedings of the IEEE SiPS*, 2021.

[11] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, “The SpiNNaker project,” *Proceedings of the IEEE*, vol. 102, no. 5, pp. 652–665, 2014.

[12] C. Mayr, S. Hoeppner, and S. Furber, “SpiNNaker 2: A 10 million core processor system for brain simulation and machine learning,” *arXiv preprint arXiv:1911.02385*, 2019.

[13] BrainChip Inc., *Akida AKD1000 Neuromorphic Processor Architecture*. BrainChip Technical Reference, 2021.

[14] G.-Q. Bi and M.-M. Poo, “Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type,” *Journal of Neuroscience*, vol. 18, no. 24, pp. 10464–10472, 1998.

[15] E. O. Neftci, H. Mostafa, and F. Zenke, “Surrogate gradient learning in spiking neural networks,” *IEEE Signal Processing Magazine*, vol. 36, no. 6, pp. 51–63, 2019.

[16] B. Rueckauer, I.-A. Lungu, Y. Hu, M. Pfeiffer, and S.-C. Liu, “Conversion of continuous-valued deep networks to efficient event-driven networks for image classification,” *Frontiers in Neuroscience*, vol. 11, 2017.

[17] Intel Labs, “Lava: A software framework for neuromorphic computing,” GitHub repository, 2022.

[18] J. K. Eshraghian, M. Ward, E. O. Neftci, X. Wang, G. Lenz, G. Dwivedi, *et al*., “Training spiking neural networks using lessons from deep learning,” *Proceedings of the IEEE*, vol. 111, no. 9, pp. 1016–1054, 2023.

[19] A. Amir, B. Taba, D. Berg, T. Melano, J. McKinstry, C. Di Nolfo, *et al*., “A low power, fully event-based gesture recognition system,” in *Proceedings of the CVPR*, 2017.

[20] N. Imam and T. A. Cleland, "Rapid online learning and robust recall in a neuromorphic olfactory circuit," *Nature Machine Intelligence*, vol. 2, pp. 181–191, 2020.
[21] A. F. Vincent, J. Larroque, N. Locatelli, N. Ben Romdhane, O. Bichler, C. Gamrat, *et al*., "Spin-transfer torque magnetic memory as a stochastic memristive synapse," in *Proceedings of the ISCAS*, 2014.
[22] M. Davies, "Benchmarks for progress in neuromorphic computing," *Nature Machine Intelligence*, vol. 1, pp. 386–388, 2019.