International IT Journal of Research (IITJR), ISSN: 3007-6706
Volume 4, Issue 1, January-March 2026
Available online at: https://itjournal.org/index.php

Neuromorphic Spiking Neural Networks for Ultra-Low-Power Event-Based Computer Vision in Autonomous Robotics

Dr. Renjiro Sato

Associate Professor, Department of Information and Communication Engineering, The University of Tokyo, Japan

ABSTRACT

The rapid advancement of autonomous robotics has increased the demand for intelligent vision systems capable of processing dynamic visual information with high speed, accuracy, and energy efficiency. Conventional computer vision architectures primarily rely on frame-based image processing and computationally intensive deep neural networks, which often result in significant power consumption, processing delays, and limited suitability for resource-constrained robotic platforms. Neuromorphic computing, particularly Spiking Neural Networks (SNNs), offers a promising alternative by emulating the information-processing mechanisms of biological neural systems. This research explores the application of neuromorphic Spiking Neural Networks for ultra-low-power event-based computer vision in autonomous robotics. The proposed approach integrates event-based vision sensors with biologically inspired spiking architectures to enable asynchronous visual information processing, real-time object recognition, motion detection, obstacle avoidance, and intelligent navigation. Unlike conventional cameras that generate continuous image frames, event-based sensors capture changes in pixel intensity independently, significantly reducing redundant data generation and computational overhead.

The proposed framework investigates the integration of leaky integrate-and-fire neurons, temporal spike encoding, synaptic learning mechanisms, and neuromorphic hardware accelerators to improve computational efficiency and response time. Furthermore, the study examines the potential of hybrid SNN architectures combining temporal feature extraction and lightweight classification mechanisms for reliable robotic perception under dynamic environmental conditions. Performance evaluation focuses on energy consumption, inference latency, recognition accuracy, event-processing efficiency, and real-time responsiveness. Comparative analysis with conventional convolutional neural networks and traditional frame-based vision systems highlights the potential advantages of neuromorphic approaches in energy-constrained autonomous robotic applications. The anticipated outcomes demonstrate that event-driven SNN architectures can substantially reduce unnecessary computations while maintaining effective visual perception and rapid decision-making capabilities. The research also addresses challenges associated with spike-based training, hardware scalability, environmental robustness, and deployment complexity. Overall, neuromorphic spiking neural networks represent a promising foundation for developing energy-efficient, adaptive, and intelligent autonomous robotic systems capable of operating in real-time environments with limited computational resources.

Keywords: Neuromorphic Computing, Spiking Neural Networks (SNNs), Event-Based Computer Vision, Ultra-Low-Power Processing, Autonomous Robotics

INTRODUCTION

The rapid development of autonomous robotics, artificial intelligence, and intelligent sensing technologies has significantly transformed the capabilities of modern robotic systems. Autonomous robots are increasingly deployed in industrial automation, healthcare, agriculture, transportation, surveillance, space exploration, and other dynamic environments. These applications require robots to perceive their surroundings, recognize objects, detect motion, identify obstacles, and make intelligent decisions with minimal human intervention. Computer vision plays a fundamental role in enabling these capabilities by providing visual information about the surrounding environment. However, conventional computer vision systems often require substantial computational resources, memory capacity, and electrical power, creating significant challenges for autonomous robots operating under strict energy and hardware constraints.

Traditional vision systems generally rely on frame-based cameras that capture images at predefined intervals. These cameras continuously generate complete image frames, even when only a small portion of the visual environment changes. Consequently, considerable computational effort is required to process redundant visual information. Conventional deep learning architectures, particularly Convolutional Neural Networks (CNNs) and Vision Transformers, further increase computational complexity through intensive matrix operations and repeated feature extraction. Although these techniques have achieved impressive results in object detection, image classification, and

visual recognition, their high energy consumption and processing requirements can limit their practical implementation in compact autonomous robotic platforms.

Neuromorphic computing has emerged as an innovative computational paradigm inspired by the structure and functioning of biological neural systems. Unlike conventional computing architectures that process information through continuous numerical operations, neuromorphic systems employ neuron-like computational elements and synaptic connections to process information through discrete electrical spikes. This approach provides opportunities for event-driven computation, parallel information processing, temporal learning, and reduced energy consumption. Spiking Neural Networks (SNNs) represent an important component of neuromorphic computing because they communicate information through sequences of discrete spikes rather than continuous-valued activations.

PROPOSED MODELS AND METHODOLOGIES

The proposed research methodology focuses on developing an energy-efficient neuromorphic computer vision framework using Spiking Neural Networks (SNNs) and event-based visual sensors for autonomous robotic applications. The framework is designed to process dynamic visual information through asynchronous event-driven mechanisms, enabling rapid perception, low-latency inference, and reduced power consumption. The proposed architecture integrates event-based data acquisition, spike encoding, temporal feature extraction, spiking neural processing, object recognition, and intelligent robotic decision-making into a unified computational pipeline.

The methodology emphasizes the optimization of neural network architecture, spike-based learning mechanisms, and hardware-level execution to achieve efficient visual perception under limited computational and energy resources.

2.1 Overall Proposed System Architecture

The proposed system consists of six major processing layers:

1. Event-based visual data acquisition.
2. Event preprocessing and spike encoding.
3. Neuromorphic feature extraction.
4. Spiking Neural Network-based classification.
5. Robotic decision-making and control.
6. Energy and performance optimization.

The event-based camera captures changes in environmental brightness and converts them into asynchronous events. These events are processed by an encoding module that transforms visual information into spike sequences. The encoded information is subsequently passed to an SNN architecture, where temporal and spatial characteristics are extracted for object recognition and motion analysis. The final classification outputs are integrated into the robotic control system to support obstacle avoidance, navigation, and real-time decision-making.

2.2 Event-Based Visual Data Acquisition Model

The first stage involves collecting visual information using Dynamic Vision Sensors (DVS) or other event-based cameras. Unlike conventional image sensors, event-based cameras generate events independently for individual pixels whenever a significant change in brightness is detected.

Each event is represented as:

$$ e_i=(x_i,y_i,t_i,p_i) $$

Where:

- x_i = Horizontal pixel coordinate.
- y_i = Vertical pixel coordinate.
- t_i = Event timestamp.
- p_i = Event polarity indicating brightness increase or decrease.

The event-generation mechanism can be represented as:

$$ \Delta L(x,y,t) \geq C $$

Where:

- ΔL represents the change in logarithmic brightness.
- C represents the predefined contrast threshold.

This mechanism enables the camera to capture dynamic visual changes without continuously generating complete image frames.

The collected event streams provide high temporal resolution and reduce redundant visual information, making them suitable for high-speed robotic perception.

RESULTS & ANALYSIS

The results and analysis section evaluates the performance of the proposed Neuromorphic Spiking Neural Network (SNN) framework for ultra-low-power event-based computer vision in autonomous robotics. The evaluation focuses on the effectiveness of event-driven visual processing, temporal feature extraction, object recognition, motion detection, obstacle avoidance, and energy-efficient computation. The proposed architecture is compared with conventional Convolutional Neural Networks (CNNs) and other event-based vision approaches to examine its potential advantages in real-time robotic applications.

The experimental evaluation considers several important performance indicators, including recognition accuracy, inference latency, power consumption, energy per inference, event-processing efficiency, and navigation success rate.

Note: The following results are illustrative research values designed to demonstrate a possible experimental evaluation. They are not measurements from an actually conducted experiment and should be replaced with validated experimental observations before publication.

3.1 Experimental Performance Evaluation

The proposed SNN architecture is evaluated against conventional frame-based CNN models and event-based deep learning architectures. The experimental framework considers dynamic environments containing moving objects, static obstacles, changing illumination conditions, and different robotic navigation scenarios.

The evaluation considers the following configurations:

- **Model 1:** Conventional CNN-based frame processing.
- **Model 2:** Event-based CNN architecture.
- **Model 3:** Conventional Spiking Neural Network.
- **Model 4:** Proposed optimized neuromorphic SNN architecture.

3.2 Object Recognition Accuracy

Object recognition accuracy represents the ability of the proposed architecture to correctly identify environmental objects using event-based visual information.

The illustrative evaluation indicates that the optimized SNN architecture can achieve competitive recognition performance while requiring fewer computational operations than conventional frame-based processing systems.

Model	Recognition accuracy (%)
Conventional CNN	94.2
Event-Based CNN	92.8
Conventional SNN	89.6
Proposed Optimized SNN	93.5

The proposed architecture achieves an illustrative recognition accuracy of 93.5%, demonstrating its potential to recognize dynamic environmental objects effectively. Although the conventional CNN produces a slightly higher accuracy, the optimized SNN offers a potentially more favorable balance between recognition performance and energy efficiency.

Comparative Analysis in Tabular Form

The comparative analysis evaluates the proposed Neuromorphic Spiking Neural Network (SNN) architecture against conventional Convolutional Neural Networks (CNNs), event-based CNNs, and conventional SNN models for ultra-low-power event-based computer vision in autonomous robotics. The comparison considers recognition accuracy, inference latency, power consumption, energy efficiency, event-processing capability, and suitability for real-time robotic applications.

4.1 Comparative Performance Evaluation

Performance Parameters	Conventional CNN	Event-Based CNN	Conventional SNN	Proposed Optimized SNN
Recognition Accuracy (%)	94.2	92.8	89.6	93.5
Inference Latency (ms)	28.5	19.2	14.8	8.6
Power Consumption (W)	8.5	5.2	2.8	1.6
Energy per Inference (mJ)	242.25	99.84	41.44	13.76
Temporal Processing	Limited	Moderate	High	Very High
Event-Driven Computation	No	Partial	Yes	Yes
Sparse Neural Activation	No	No	Yes	Optimized
Real-Time Responsiveness	Moderate	High	High	Very High
Embedded Deployment	Hardware Intensive	Moderate	Suitable	Highly Suitable
Energy Efficiency	Low	Moderate	High	Very High

Note: *Numerical values are illustrative research estimates, not experimentally validated measurements. Energy values are calculated from the corresponding power and latency figures.*

4.2 Comparison of Architectural Characteristics

Architectural Features	CNN Architecture	Event-Based CNN	Conventional SNN	Proposed SNN
Input Representation	Image Frames	Event Frames/Voxel Grids	Spike Sequences	Asynchronous Event Spikes
Processing Mechanism	Continuous Numerical Operations	Event-Derived Processing	Spike-Based Processing	Optimized Event-Driven Processing
Neuron Model	Artificial Neurons	Artificial Neurons	LIF Neurons	Optimized LIF Neurons
Temporal Information	Limited	Moderate	Strong	Strong
Learning Mechanism	Backpropagation	Backpropagation	STDP/Surrogate Gradients	Hybrid Spike-Based Learning
Data Redundancy	High	Reduced	Low	Minimized
Hardware Requirements	GPU/CPU	GPU/CPU/Accelerator	Neuromorphic Hardware	Optimized Neuromorphic Hardware
Primary Application	General Computer Vision	Dynamic Vision	Temporal Recognition	Autonomous Robotic Perception

4.3 Comparative Analysis of Robotic Applications

Application Area	Conventional CNN	Event-Based CNN	Proposed SNN
Object Recognition	High accuracy	High accuracy	Competitive accuracy
Motion Detection	Frame-dependent	Efficient	Highly suitable
Obstacle Avoidance	Computationally intensive	Faster event processing	Low-latency perception
Autonomous Navigation	High processing demand	Improved temporal sensing	Energy-efficient perception
Low-Light Operation	Depends on image quality	Suitable for dynamic changes	Suitable with robust event encoding
Rapid Movement Detection	Potential motion blur	High temporal resolution	High temporal responsiveness
Battery-Powered Robotics	High energy demand	Moderate energy demand	Potentially very low energy demand

LIMITATIONS & DRAWBACKS

Although Neuromorphic Spiking Neural Networks (SNNs) offer significant potential for ultra-low-power event-based computer vision in autonomous robotics, several technical, computational, and practical limitations affect their implementation and large-scale deployment. Despite their advantages in asynchronous processing, temporal learning, and energy-efficient computation, SNN-based vision systems still face challenges related to training complexity, recognition accuracy, hardware availability, sensor reliability, and environmental adaptability.

The major limitations and drawbacks of the proposed framework are discussed below.

5.1 Complexity of SNN Training

One of the major challenges associated with Spiking Neural Networks is their training complexity. Unlike conventional artificial neural networks, SNNs generate discrete spikes through threshold-based activation mechanisms. These discontinuous operations make direct gradient-based optimization difficult.

Although surrogate-gradient learning and Spike-Timing-Dependent Plasticity (STDP) provide alternative training approaches, achieving stable convergence and optimal performance remains challenging.

5.2 Limited Recognition Accuracy

Conventional CNN architectures often achieve higher recognition accuracy in complex visual classification tasks because of their mature feature extraction and optimization mechanisms. SNN architectures may experience reduced accuracy when processing highly complex objects, overlapping visual patterns, or environmental scenes containing multiple moving elements.

Maintaining competitive accuracy while minimizing neuron activation and computational energy remains an important research challenge.

5.3 Event-Based Sensor Limitations

Event-based cameras offer high temporal resolution and reduced redundant data generation. However, their performance depends on environmental brightness changes and sensor characteristics.
Important limitations include:

- Difficulty detecting stationary objects without sufficient brightness changes.
- Background event noise under challenging illumination conditions.
- Sensitivity to sensor threshold configurations.
- Limited availability of conventional image information.
- Challenges in interpreting scenes with minimal visual activity.

These limitations may affect object recognition and environmental understanding in autonomous robotic applications.

5.4 Hardware Compatibility and Availability

Neuromorphic SNN implementations often require specialized hardware or optimized computational architectures. Although neuromorphic processors offer promising energy-efficient processing capabilities, their availability and development ecosystems remain more limited than those of conventional CPUs and GPUs.
Furthermore, differences in supported neuron models, programming interfaces, memory architectures, and spike communication mechanisms can create hardware compatibility challenges.

5.5 Difficulties in Real-Time Deployment

Although SNNs are designed for event-driven processing, achieving consistently low latency in practical robotic environments is not guaranteed. Data preprocessing, event synchronization, network inference, communication overhead, and robotic control operations may introduce additional delays.
Therefore, optimizing the entire perception-to-action pipeline is essential for reliable real-time operation.

CONCLUSION

The research on **"Neuromorphic Spiking Neural Networks for Ultra-Low-Power Event-Based Computer Vision in Autonomous Robotics"** highlights the potential of biologically inspired computing architectures for developing energy-efficient, intelligent, and responsive autonomous robotic systems. The integration of Spiking Neural Networks (SNNs) with event-based vision sensors provides an alternative to conventional frame-based computer vision architectures, which often require substantial computational resources and electrical power. By utilizing asynchronous visual information, temporal spike encoding, and sparse neural activation, neuromorphic systems can reduce redundant computations while supporting rapid environmental perception.

The proposed framework integrates event-based data acquisition, event preprocessing, spike encoding, Leaky Integrate-and-Fire (LIF) neuron models, temporal feature extraction, and SNN-based object recognition. It further investigates Spike-Timing-Dependent Plasticity (STDP), surrogate-gradient learning, and hardware-aware optimization to improve neural processing efficiency. The integration of these mechanisms provides a foundation for intelligent robotic applications involving object recognition, motion detection, obstacle avoidance, and autonomous navigation.

The illustrative performance analysis indicates that the proposed optimized SNN architecture has the potential to achieve competitive recognition accuracy while reducing inference latency and energy consumption. Compared with the conventional CNN baseline, the illustrative results demonstrate approximately 69.8% lower inference latency, 81.2% lower power consumption, and 94.3% lower energy consumption per inference. These figures represent hypothetical research estimates and require experimental verification using actual event-based datasets and neuromorphic hardware.

Furthermore, the research identifies important implementation challenges, including SNN training complexity, event-sensor limitations, hardware compatibility, environmental adaptability, network scalability, and safety requirements. Addressing these challenges is essential for establishing reliable neuromorphic vision systems capable of operating in complex real-world robotic environments.

Future research should concentrate on developing advanced spike-based learning algorithms, adaptive event encoding techniques, hybrid SNN architectures, and optimized neuromorphic hardware implementations. The integration of event-based vision with complementary sensing technologies, including LiDAR, radar, and conventional cameras, may further improve perception reliability and environmental understanding. Additional investigations into real-world robotic deployment, full-system energy measurement, and standardized benchmarking will be necessary to validate the practical advantages of the proposed framework.

REFERENCES

1. Amir, A., Taba, B., Berg, D., Melano, T., McKinstry, J., Di Nolfo, C., Nayak, T. K., Andreopoulos, A., Garreau, G., Mendoza, M., Kusnitz, J., Debole, M., Esser, S. K., Delbruck, T., Flickner, M. D., & Modha, D. S. (2017). A low power, fully event-based gesture recognition system. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7388–7397.
2. Brandli, C., Berner, R., Yang, M., Liu, S.-C., & Delbruck, T. (2014). A 240 × 180 130 dB 3 µs latency global shutter spatiotemporal vision sensor. *IEEE Journal of Solid-State Circuits*, *49*(10), 2333–2341.
3. Chicca, E., Stefanini, F., Bartolozzi, C., & Indiveri, G. (2014). Neuromorphic electronic circuits for building autonomous cognitive systems. *Proceedings of the IEEE*, *102*(9), 1367–1388.
4. Davies, M., Srinivasa, N., Lin, T.-H., Chinya, G., Cao, Y., Choday, S. H., Dimou, G., Joshi, P., Imam, N., Jain, S., Liao, Y., Lin, C.-K., Lines, A., Liu, R., Mathaikutty, D., McCoy, S., Paul, A., Tse, J., Venkataramanan, G., ... Wang, H. (2018). Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, *38*(1), 82–99.
5. Delbruck, T., & Lichtsteiner, P. (2007). Fast sensory motor control based on event-based hybrid neuromorphic-procedural system. *Proceedings of the 2007 IEEE International Symposium on Circuits and Systems*, 845–848.
6. Furber, S. B., Galluppi, F., Temple, S., & Plana, L. A. (2014). The SpiNNaker project. *Proceedings of the IEEE*, *102*(5), 652–665.
7. Gallego, G., Delbruck, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A. J., Conradt, J., Daniilidis, K., & Scaramuzza, D. (2022). Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(1), 154–180. https://doi.org/10.1109/TPAMI.2020.3008413
8. Gehrig, D., Rebecq, H., Gallego, G., & Scaramuzza, D. (2019). End-to-end learning of representations for asynchronous event-based data. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 5633–5643.
9. Gerstner, W., & Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge University Press.
10. Hendy, H., & Merkel, C. (2022). Review of spike-based neuromorphic computing for brain-inspired vision: Biology, algorithms, and hardware. *Journal of Electronic Imaging*, *31*(1), 010901. https://doi.org/10.1117/1.JEI.31.1.010901
11. Indiveri, G., Linares-Barranco, B., Hamilton, T. J., Van Schaik, A., Etienne-Cummings, R., Delbruck, T., Liu, S.-C., Dudek, P., Häfliger, P., Renaud, S., Schemmel, J., Cauwenberghs, G., Arthur, J., Hynna, K., Folowosele, F., SAÏGHI, S., Serrano-Gotarredona, T., Wijekoon, J., & Wang, Y. (2011). Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience*, *5*, Article 73.
12. Lichtsteiner, P., Posch, C., & Delbruck, T. (2008). A 128 × 128 120 dB 15 µs latency asynchronous temporal contrast vision sensor. *IEEE Journal of Solid-State Circuits*, *43*(2), 566–576.
13. Maass, W. (1997). Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, *10*(9), 1659–1671.

14. Merolla, P. A., Arthur, J. V., Alvarez-Icaza, R., Cassidy, A. S., Sawada, J., Akopyan, F., Jackson, B. L., Imam, N., Guo, C., Nakamura, Y., Brezzo, B., Vo, I., Esser, S. K., Appuswamy, R., Taba, B., Amir, A., Flickner, M. D., Risk, W. P., Manohar, R., & Modha, D. S. (2014). A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science, 345*(6197), 668–673.
15. Mueggler, E., Rebecq, H., Gallego, G., Delbruck, T., & Scaramuzza, D. (2017). The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and SLAM. *The International Journal of Robotics Research, 36*(2), 142–149.
16. Neftci, E. O., Mostafa, H., & Zenke, F. (2019). Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Processing Magazine, 36*(6), 51–63.
17. Orchard, G., Meyer, C., Etienne-Cummings, R., Posch, C., Thakor, N., & Benosman, R. (2015). HFirst: A temporal approach to object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(10), 2028–2040.
18. Pfeiffer, M., & Pfeil, T. (2018). Deep learning with spiking neurons: Opportunities and challenges. *Frontiers in Neuroscience, 12*, Article 774.
19. Rebecq, H., Gehrig, D., & Scaramuzza, D. (2018). ESIM: An open event camera simulator. *Proceedings of the Conference on Robot Learning*, 969–982.
20. Rueckauer, B., Lungu, I.-A., Hu, Y., Pfeiffer, M., & Liu, S.-C. (2017). Conversion of continuous-valued deep networks to efficient event-driven networks for image classification. *Frontiers in Neuroscience, 11*, Article 682.