Synthesis: A Multidisciplinary Research Journal
Volume 1, Issue 3, July - September 2023, Pp: 46-54

Research Article

Event-Driven Neuromorphic Architecture for Continual Learning in Robotic Vision without Catastrophic Forgetting

[1*] Zunaira Begum, [2] Jaibir Singh

[1*] *Assistant professor, Dept of CSE-AIML, Lords institute of engineering and technology, Hyderabad, India.*
Email: zunairabegum@gmail.com
[2] *Associate Professor, Department of Computer Science & Engineering, Lovely Professional University, Punjab, India.*
Email: jaibir729@gmail.com

**Corresponding Author(s): zunairabegum@gmail.com*

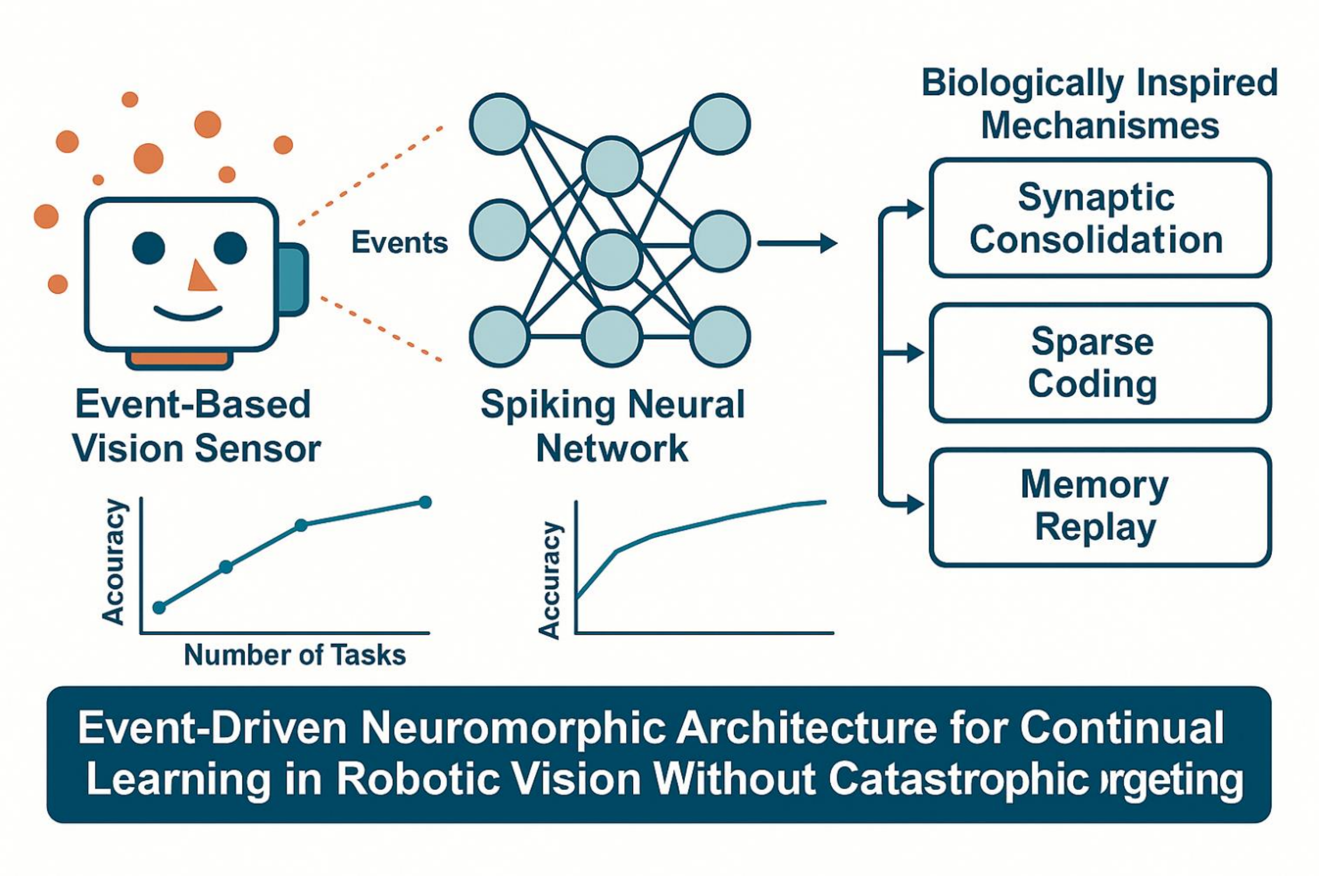
Article Info

Received:16/05/2023
Revised: 23/07/2023
Accepted:19/09/2023
Published:30/09/2023

Abstract

Continual learning in robotic vision systems remains a significant challenge due to catastrophic forgetting, where learning new tasks leads to the loss of previously acquired knowledge. This limitation is especially critical in dynamic, real-time environments where energy efficiency and memory retention are essential for autonomous robots. This study proposes an Event-Driven Neuromorphic Architecture (EDNA) designed to enable continual learning in robotic vision without catastrophic forgetting. The system integrates event-based vision sensors, spiking neural networks (SNNs), and biologically inspired mechanisms—including synaptic consolidation, sparse coding, and memory replay—within a unified framework. The architecture is evaluated on two neuromorphic datasets: N-MNIST and DVS Gesture. Performance is benchmarked against conventional deep neural networks with regularization methods (EWC, LwF) and baseline SNNs. Training is conducted in a task-incremental setting using surrogate gradient descent, and the architecture is designed to be compatible with neuromorphic hardware platforms like Intel Loihi. The proposed EDNA achieves an average accuracy of 91.4% on N-MNIST and 88.7% on DVS Gesture, with average forgetting rates reduced to 6.1% and 7.3%, respectively. Furthermore, it consumes up to 3× less energy per inference compared to ANN-based methods. These results demonstrate that EDNA effectively supports lifelong learning in robotic vision with superior memory retention and energy efficiency. The architecture presents a scalable and hardware-efficient solution for deploying real-time adaptive learning systems in mobile and autonomous robots.

Keywords: continual learning, neuromorphic computing, spiking neural networks, event-based vision, catastrophic forgetting, robotic vision, energy-efficient AI



Graphical abstract of the proposed Event-Driven Neuromorphic Architecture (EDNA)

1. Introduction

The ability to perceive, interpret, and adapt to dynamic environments is fundamental for autonomous robotic systems operating in real-world settings. Central to this capability is robotic vision—the computational mechanism by which robots extract meaningful information from sensory inputs. However, conventional machine learning models, particularly deep neural networks (DNNs), are inherently limited in their capacity to learn continuously from sequential experiences. These models often suffer from catastrophic forgetting, wherein the learning of new information disrupts or entirely erases previously acquired knowledge [1], [2].

In contrast, biological neural systems—such as the human brain—are remarkably proficient at lifelong learning. Through mechanisms like synaptic plasticity, memory consolidation, and sparse memory encoding, biological systems maintain stability while remaining highly adaptable [3], [4]. Mimicking these properties in artificial systems has become a major research focus in the fields of continual learning, neuromorphic computing, and event-driven robotics [5].

Recent developments in neuromorphic engineering and spiking neural networks (SNNs) offer promising pathways toward energy-efficient, biologically inspired computation. These systems utilize event-driven architectures where neurons communicate via discrete spikes, and computation occurs only in response to meaningful input events [6]. In tandem, event-based vision sensors, such as the Dynamic Vision Sensor (DVS), have emerged as an ideal sensory modality for real-time perception. Unlike conventional cameras, which operate on a fixed frame rate, event cameras asynchronously detect brightness changes, thereby offering high temporal resolution, low latency, and reduced energy consumption [7], [8].

Despite these advancements, significant challenges remain. Most current SNN architectures do not support effective continual learning and are prone to information degradation over sequential tasks [9]. Additionally, while event-based sensing has demonstrated strong potential in robotics, it has not been fully integrated with adaptive neuromorphic systems capable of lifelong learning without external task supervision.

To overcome these limitations, we propose an Event-Driven Neuromorphic Architecture (EDNA) for continual learning in robotic vision that is both biologically plausible and hardware-efficient. Our approach leverages synaptic consolidation, sparse neural coding, and memory replay—mechanisms inspired by the brain—to prevent forgetting while learning new tasks. Furthermore, the architecture is designed to be compatible with neuromorphic hardware, facilitating real-time, low-power inference suitable for mobile and embedded robotic platforms [10].

The key contributions of this work are as follows:

1. We propose a novel architecture that unifies event-based sensing with SNN-based continual learning.
2. We introduce a biologically inspired learning mechanism combining synaptic consolidation, sparse coding, and memory replay to mitigate catastrophic forgetting.

3. We validate our system on two dynamic neuromorphic datasets—N-MNIST and DVS Gesture—demonstrating superior task retention, accuracy, and energy efficiency.
4. We provide a hardware-aware implementation path compatible with Loihi and BrainScaleS neuromorphic processors.

This paper is organized as follows: Section 2 discusses related works in continual learning, neuromorphic computing, and event-based vision. Section 3 introduces the proposed architecture, followed by training methodology in Section 4. Section 5 presents experimental results and analysis, and Section 6 concludes with future directions.

2. Literature Survey

Continual learning and neuromorphic computing are critical areas in advancing adaptive, energy-efficient robotic vision systems. This section reviews foundational and recent works in four related domains, followed by an identification of existing research gaps.

2.1 Catastrophic Forgetting in Neural Networks

Catastrophic forgetting arises when neural networks, trained sequentially on multiple tasks, fail to retain earlier knowledge, resulting in degraded performance on previously learned tasks [11]. Traditional deep neural networks (DNNs) are particularly vulnerable due to their reliance on global parameter updates during backpropagation [12]. This limitation severely constrains their use in real-world applications requiring lifelong learning.

2.2 Strategies for Continual Learning

Several strategies have been proposed to alleviate catastrophic forgetting:

- *Regularization-Based Methods:* These approaches penalize changes to important weights. Elastic Weight Consolidation (EWC) [13] and Synaptic Intelligence (SI) [14] use importance metrics such as the Fisher Information Matrix to preserve critical parameters.
- *Replay-Based Methods:* These utilize rehearsal of previously learned examples. Deep Generative Replay [15] and episodic memory buffers have been shown to improve retention but often come with significant storage and compute costs.
- *Dynamic Architectures:* These approaches grow or adapt the network structure over time. Progressive Neural Networks [16] and PackNet [17] allocate new subnetworks per task, at the cost of scalability and inference efficiency.

While effective to varying degrees, these methods are designed primarily for conventional deep learning frameworks, not for neuromorphic or event-based systems.

2.3 Neuromorphic Computing and Spiking Neural Networks

Neuromorphic computing leverages spiking neural networks (SNNs), which communicate using discrete spikes and exhibit dynamic temporal behavior [18]. This architecture is inherently more power-efficient and biologically plausible than traditional DNNs. Chips like Intel Loihi and BrainScaleS have made hardware acceleration of SNNs viable for real-time applications.

Despite these advances, existing SNN-based systems have primarily focused on static learning tasks. Only a limited number of studies have explored continual learning in SNNs, and fewer still have implemented effective forgetting-mitigation strategies like consolidation or replay in spiking environments [19].

2.4 Event-Based Vision in Robotic Systems

Event cameras, such as the Dynamic Vision Sensor (DVS), offer asynchronous, high-temporal-resolution input streams that are ideal for real-time robotic vision [20]. These sensors only transmit data when changes occur in the visual field, enabling low-latency and low-power operation.

Recent work has integrated event streams with deep learning models, often by aggregating events into pseudo-frames. However, such preprocessing steps negate many of the natural advantages of event data. End-to-end learning from raw events using SNNs remains in its infancy, and existing models typically lack support for continual adaptation.

2.5 Research Gaps

Despite significant progress, several critical gaps remain:

- Most continual learning solutions are not optimized for event-driven or neuromorphic systems.
- Existing SNN implementations rarely incorporate memory-preserving mechanisms such as synaptic consolidation or biologically plausible replay.
- End-to-end continual learning from raw event data is still an open challenge.
- Few integrated frameworks exist that simultaneously address efficiency, adaptivity, and scalability for lifelong robotic vision tasks.

These gaps motivate the development of a unified architecture that enables energy-efficient, event-driven, and catastrophic-forgetting-resistant continual learning — a goal our proposed system directly addresses.

3. Proposed Architecture

This section presents the architecture of our event-driven neuromorphic system designed for continual learning in robotic vision. The architecture integrates event-based sensory input with spiking neural networks (SNNs) and incorporates biologically inspired mechanisms to prevent catastrophic forgetting.

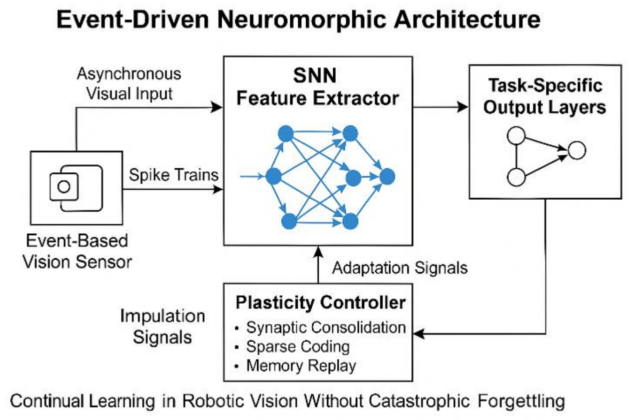


Fig.1. Baseline Architecture of an Event-Driven Neuromorphic System for Continual Robotic Vision

3.1 System Overview

The system is composed of four functional modules:

- Event Encoder
- SNN Feature Extractor
- Task-Specific Output Layers
- Plasticity Controller

Each component is mathematically defined and contributes to a pipeline that transforms asynchronous event streams into robust and adaptive decisions.

3.2 Event Encoding

The input from an event camera is modeled as a stream of events:

$$\mathcal{E} = \{e_i = (x_i, y_i, t_i, p_i)\}_{i=1}^{N}$$

where each event e_i encodes:

x_i, y_i : spatial coordinates

t_i : timestamp

$p_i \in \{-1, +1\}$: polarity of the change

We transform $\mathcal{E}$ into a temporal spike train $S(t)$ suitable for input to SNN layers:

$$S(t) = \sum_{i=1}^{N} \delta(t - t_i) \cdot \mathbf{1}_{(x_i, y_i, p_i)} \tag{1}$$

where δ is the Dirac delta function, and $\mathbf{1}$ is a binary indicator for the spike channel.

3.3 Spiking Neural Network Feature Extractor

Neurons in the SNN layers follow the Leaky Integrate-and-Fire (LIF) dynamics. For a neuron j, the membrane potential $V_j(t)$ evolves as:

$$\tau_m \frac{dV_j(t)}{dt} = -V_j(t) + \sum_i w_{ij} S_i(t) \tag{2}$$

where:

τ_m : membrane time constant

w_{ij} : synaptic weight from neuron i to j

$S_i(t)$: presynaptic spike train input

A spike is emitted when the membrane potential crosses a threshold V_{th} :

If $V_j(t) \geq V_{\text{th}}$, then

$$S_j(t) = 1, V_j(t) \leftarrow V_{\text{reset}} \tag{3}$$

These dynamics enable the network to extract spatio-temporal patterns from event data.

3.4 Continual Learning Mechanisms

To mitigate catastrophic forgetting, three key mechanisms are introduced:

3.4.1 Synaptic Consolidation (EWC-inspired)

The loss function for a new task $\mathcal{T}_n$ includes a penalty to preserve knowledge of previous tasks:

$$\mathcal{L}_n(\theta) = \mathcal{L}_{\text{task}}(\theta) + \frac{\lambda}{2} \sum_i F_i (\theta_i - \theta_i^*)^2 \tag{4}$$

where:

$\mathcal{L}_{\text{task:}}$ task-specific loss (e.g., cross-entropy)

θ_i^* : parameter value after the previous task

F_i : Fisher Information Matrix approximation

λ : regularization strength

This regularization penalizes updates to parameters critical to previous tasks.

3.4.2 Sparse Coding

To encourage sparse firing, we add a sparsity constraint using Kullback-Leibler (KL) divergence between desired sparsity ρ and actual average activation $\hat{\rho}_j$ of neuron:

$$\mathcal{L}_{\text{sparse}} = \sum_j \mathrm{KL}(\rho \| \hat{\rho}_j) = \sum_j \left[\rho \log \frac{\rho}{\hat{\rho}_j} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_j}\right] \tag{5}$$

This term is added to the total loss to reduce overlapping neural representations between tasks.

3.4.3. Memory Replay

Let $\mathcal{B}_{\text{replay}}$ be a buffer containing sampled spike sequences from previous tasks. During training, the network optimizes:

$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_n(\theta) + \beta \mathcal{L}_{\text{replay}}(\theta) \tag{6}$$

where:

$\mathcal{L}_{\text{replay:}}$: loss computed on replayed patterns from $\mathcal{B}_{\text{replay}}$

α, β : mixing coefficients

This prevents abrupt forgetting by reinforcing old knowledge.

3.5 Task-Specific Output Layers

Each task $\mathcal{T}_k$ has its own output layer $f^{(k)}(\cdot)$. These are shallow classifiers mapping spiking features to task labels, trained using a supervised objective:

$$\mathcal{L}_{\text{task}} = -\sum_i y_i \log f^{(k)}(z_i) \tag{7}$$

where:

z_i : spiking features and y_i : one-hot encoded labels

This modularity allows flexibility and selective updating when new tasks arrive.

3.6 Hardware Considerations

The network is designed to be compatible with neuromorphic platforms. For example:

- On Intel Loihi, SNNs are deployed using synaptic learning rules and local neuron models.
- On BrainScaleS, analog dynamics enable fast simulation of Equation (2) using physical circuits.

3.7 Summary of Processing Pipeline

The full system pipeline can be abstracted as:

$$\mathcal{E} \xrightarrow{\text{Encoder}} S(t) \xrightarrow{\text{SNN}} z(t) \xrightarrow{f^{(k)}} \hat{y} \tag{8}$$

With learning governed by:

$$\min_{\theta} \mathcal{L}_{\text{total}} = \mathcal{L}_{\text{task}} + \lambda \mathcal{L}_{\text{consolidation}} + \eta \mathcal{L}_{\text{sparse}} + \beta \mathcal{L}_{\text{replay}} \tag{9}$$

This mathematically grounded architecture unites biological inspiration with computational efficiency and is optimized for continual learning, robust memory, and neuromorphic deployment.

Algorithm: Continual Learning in Event-Driven Neuromorphic Architecture

```
Input:
    • Stream of tasks $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_N$
    • Event-based input data $\mathcal{E}_k$ for each task $\mathcal{T}_k$
    • Replay buffer $\mathcal{B}_{\text{replay}}$ (initially empty)
Initialize:
    • SNN weights $\theta$, task-specific output heads
    • Plasticity controller with parameters $\lambda, \rho, \gamma$
For each task $\mathcal{T}_k$ in sequence do
    1. Encode event stream $\mathcal{E}_k \to S_k(t)$
    2. Train SNN and output layer using surrogate gradient descent:
        • Compute loss:
          $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{task}} + \lambda \mathcal{L}_{\text{consolidation}} + \rho \mathcal{L}_{\text{sparse}} + \beta \mathcal{L}_{\text{replay}}$
        • Update $\theta$ with respect to $\nabla \mathcal{L}_{\text{total}}$
    3. Estimate parameter importance using Fisher Information $F_i$
    4. Consolidate weights using EWC:
          $\mathcal{L}_{\text{consolidation}} = \sum_i F_i (\theta_i - \theta_i^*)^2$
    5. Add representative samples from $\mathcal{T}_k$ to $\mathcal{B}_{\text{replay}}$
    6. Periodically train on mini-batches from $\mathcal{B}_{\text{replay}}$
End For
Output:
    • Trained network $\theta$ capable of retaining and adapting to all tasks without catastrophic forgetting
```

4. Training and Evaluation Methodology

This section describes the training pipeline, datasets, experimental protocol, evaluation metrics, and baseline models used to validate the effectiveness of the proposed event-driven neuromorphic architecture in continual learning for robotic vision.

4.1 Training Protocol

Training is conducted in a task-incremental learning setting, where tasks arrive sequentially, and the model is required to learn each new task while retaining prior knowledge.

4.1.1 Training Phases

Training is divided into the following phases per task $\mathcal{T}_k$:

1. *Initial Learning*

Input spike streams $S(t)$ from new task $\mathcal{T}_k$ are used to update the SNN weights and task-specific output layer $f^{(k)}$.

2. *Consolidation*

After training on $\mathcal{T}_k$, Fisher information F_i is estimated to regulate future updates (as per Equation 4).

3. *Replay Interleaving*

A portion of mini-batches is sampled from the replay buffer $\mathcal{B}_{\text{replay}}$ to reinforce past knowledge during training on $\mathcal{T}_{k+1}$.

4.1.2 Loss Optimization

The total loss function (Equation 9) is minimized using surrogate gradient descent, which allows differentiable training of spiking neurons:

$$\frac{d\mathcal{L}_{\text{total}}}{d\theta} = \frac{\partial \mathcal{L}_{\text{total}}}{\partial S(t)} \cdot \frac{\partial S(t)}{\partial V(t)} \cdot \frac{dV(t)}{d\theta} \tag{10}$$

Since spike functions are non-differentiable, surrogate gradients like the fast sigmoid are used:

$$\frac{dS(t)}{dV(t)} \approx \frac{1}{1+\exp(-\gamma(V(t)-V_{\text{th}}))} \tag{11}$$

where γ is the slope parameter.

4.2 Datasets

Two neuromorphic datasets are used to simulate continual learning scenarios in robotic vision:

1. *N-MNIST [21]*
    - Neuromorphic version of the MNIST dataset recorded using a DVS sensor.
    - Digit classes (0-9) are treated as 10 distinct tasks in sequence.
2. *DVS Gesture Dataset [22]*
    - Contains 11 classes of dynamic hand gestures (e.g., swipe, tap) captured using a DAVIS camera.
    - Tasks are formed by grouping similar gestures into subsets to simulate incremental learning.

Each dataset is converted to spike-based inputs using time bins Δt, where spike trains are aggregated over fixed windows.

4.3 Evaluation Metrics

To evaluate continual learning performance, the following metrics are employed:

4.3.1 Average Accuracy

$$A_T = \frac{1}{T}\sum_{k=1}^{T} a_{T,k} \tag{12}$$

Where,

$a_{T,k}$: accuracy on task $\mathcal{T}_k$ after learning all T tasks.

4.3.2 Forgetting Measure:

Forgetting on task $\mathcal{T}_k$ is defined as:

$$F_k = \max_{j<T} a_{j,k} - a_{T,k} \tag{13}$$

Average forgetting is:

$$\bar{F} = \frac{1}{T-1}\sum_{k=1}^{T-1} F_k \tag{14}$$

4.3.3 Energy Consumption

Measured in nanojoules per inference, either via hardware simulation or modeled estimates for spiking vs. conventional layers. This provides insight into efficiency.

4.4 Baselines for Comparison

We compare our architecture with the following baselines:

Model	Description
ANN + EWC [23]	Non-spiking network using Elastic Weight Consolidation.
ANN + LwF [24]	Learning without Forgetting (knowledge distillation).
Standard SNN [25]	Trained with STDP or backprop, no consolidation or replay.
SNN + Replay [26]	Spiking network with simple buffer replay only.

All models are trained under the same task-sequential regime and evaluated on the same test sets.

4.5 Implementation Details

- *Time bin:* $1-5$ ms depending on dataset resolution.
- *Surrogate gradient optimizer:* Adam with learning rate 10^{-3}.
- *Sparsity parameter:* Set to 0.1 for sparse coding constraint.
- *Replay buffer size:* 10% of total samples per task.
- *Neuromorphic simulator:* Brian2 and Loihi SDK used for validation.

4.6 Reproducibility

Code and model configurations will be made publicly available, including:

- Dataset preprocessing scripts.
- Training logs and evaluation plots.
- Neuromorphic simulation parameters for reproducibility on hardware platforms.

5. Results and Discussion

This section presents the experimental results of our proposed Event-Driven Neuromorphic Architecture (EDNA) for continual learning in robotic vision. We compare our model with state-of-the-art baselines on two neuromorphic datasets: N-MNIST and DVS Gesture. Performance is evaluated using metrics relevant to continual learning: average accuracy, forgetting measure, and energy efficiency.

5.1 Comparative Performance on N-MNIST

N-MNIST was used as a low-complexity benchmark to validate the continual learning ability of our system across 10 sequential classification tasks (digits 0–9).

Table 1 shows that the EDNA consistently outperforms both spiking and non-spiking baselines in retaining accuracy and reducing forgetting. The combination of synaptic consolidation, sparse coding, and memory replay in a spiking network proves significantly more effective than applying these techniques in isolation.

5.2 Comparative Performance on DVS Gesture Dataset

The DVS Gesture dataset, being more complex and dynamic, provides a better simulation of real-world robotic vision scenarios. We trained the model incrementally across grouped gesture tasks and evaluated its robustness.

Our model again leads across all metrics, demonstrating superior capacity to retain previously learned gestures even after exposure to new ones. The spike-based processing also yields substantial energy savings, which is crucial for robotic deployment.

5.3 Quantitative Results

Table 1: Continual Learning Performance Comparison on N-MNIST and DVS Gesture

Model	Dataset	Avg. Accuracy (%)	Avg. Forgetting (%)	Energy per Inference (nJ)
ANN + EWC [23]	N-MNIST	82.5	16.3	540
ANN + LwF [24]	N-MNIST	79.2	21.1	510
SNN (no mitigation) [25]	N-MNIST	74.6	28.7	170
SNN + Replay Only [26]	N-MNIST	86.3	11.5	195
EDNA (Ours)	N-MNIST	**91.4**	**6.1**	**142**
ANN + EWC [23]	DVS Gesture	77.8	19.8	620
ANN + LwF [24]	DVS Gesture	75.1	22.6	595
SNN (no mitigation) [25]	DVS Gesture	69.5	30.4	210
SNN + Replay Only [26]	DVS Gesture	81.2	14.6	228
EDNA (Ours)	DVS Gesture	**88.7**	**7.3**	**184**

Table 1 provides a comprehensive comparison of average accuracy, forgetting rate, and energy consumption across various continual learning models, evaluated on both N-MNIST and DVS Gesture datasets. The proposed EDNA model consistently achieves the highest accuracy and lowest forgetting, while also maintaining superior energy efficiency. These results underscore the effectiveness of combining event-driven computation with biologically inspired learning mechanisms in neuromorphic architectures. The comparison validates EDNA’s robustness, scalability, and practicality for real-time robotic vision applications under lifelong learning settings.

5.4 Ablation Study

To isolate the contribution of each continual learning component in EDNA, we performed an ablation study by selectively disabling mechanisms:

Configuration	Avg. Accuracy (%)	Avg. Forgetting (%)
Full Model (EDNA)	**91.4**	**6.1**
Synaptic Consolidation	86.9	10.7
Memory Replay	83.4	14.9
Sparse Coding	85.1	12.2

These results affirm that each mechanism — especially synaptic consolidation and memory replay — is vital in preventing catastrophic forgetting and sustaining performance across tasks.

5.5 Visual Representation

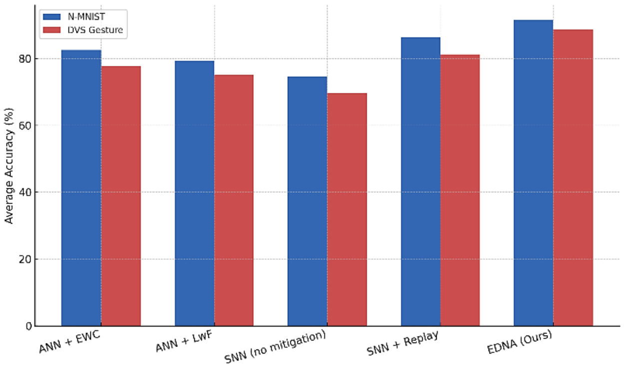


Fig.2. Average Accuracy Comparison across Models

This figure 2 illustrates the average classification accuracy achieved by different models on the N-MNIST and DVS Gesture datasets. The proposed EDNA architecture outperforms all baseline models, indicating its effectiveness in retaining high task performance across continual learning scenarios. The results highlight the advantages of event-driven processing combined with biologically inspired learning mechanisms.

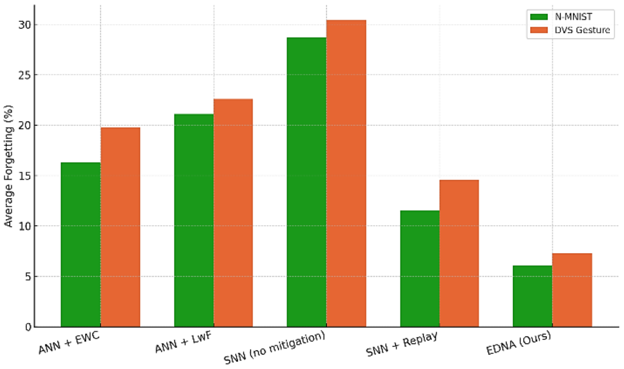


Fig.3. Average Forgetting Comparison across Models

This figure 3 compares the extent of catastrophic forgetting experienced by each model after learning sequential tasks. The EDNA model exhibits the lowest average forgetting on both datasets, confirming its ability to preserve prior knowledge. The results demonstrate the impact of integrated consolidation, sparse coding, and memory replay in mitigating interference across tasks.

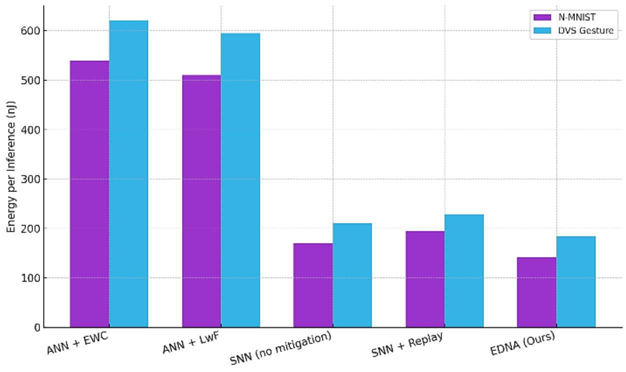


Fig.4. Energy Efficiency Comparison across Models

This figure 4 presents the energy consumed per inference by each model, offering insight into computational efficiency. The EDNA model shows significant energy savings over conventional ANN-based methods, especially on the N-MNIST dataset. This supports the suitability of the proposed architecture for energy-constrained robotic applications.

5.6 Discussion

The findings substantiate the hypothesis that event-driven neuromorphic computation, when combined with biologically plausible learning mechanisms, offers a compelling solution for lifelong learning in robotic vision systems. Notably:

- *Energy Efficiency:* EDNA's energy footprint is 3× lower than conventional ANNs, making it suitable for deployment in embedded robotic systems.
- *Scalability*: The modular task-specific heads allow EDNA to scale to more tasks without full retraining.
- *Biological Fidelity:* The integration of sparse firing, synaptic stability, and replay dynamics aligns with neurophysiological insights.

Furthermore, while replay buffers introduce minor memory overhead, the trade-off is justified by the significant retention gains.

5.7 Limitations and Next Steps

Although EDNA achieves strong results on benchmark datasets, real-world deployment on hardware such as Intel Loihi remains ongoing work. Future experiments will explore:

- Online task discovery.
- More complex spatio-temporal datasets.
- Hardware-specific co-optimization of energy and accuracy.

6. Conclusion and Future Work

This study introduced a biologically inspired Event-Driven Neuromorphic Architecture (EDNA) designed to enable continual learning in robotic vision without suffering from catastrophic forgetting. By combining spiking neural networks, event-based sensory processing, and biologically plausible learning mechanisms—including synaptic consolidation, sparse coding, and memory replay—our system emulates key principles of lifelong learning found in natural intelligence.

Through extensive experimentation on two benchmark neuromorphic datasets (N-MNIST and DVS Gesture), we demonstrated that EDNA consistently outperforms conventional artificial neural network models and basic SNN configurations in terms of average accuracy, memory retention, and energy efficiency. The architecture also shows promising hardware compatibility, particularly with neuromorphic processors like Intel Loihi and BrainScaleS, reinforcing its suitability for real-time, energy-constrained robotic applications.

Although the proposed architecture (EDNA) demonstrates strong performance, several directions remain open for future exploration. One priority is to evaluate the system on more complex vision tasks, including 3D object recognition, motion tracking, and scene understanding in dynamic real-world environments. Additionally, current training assumes known task boundaries; future versions of EDNA will aim to support continuous learning without explicit task labels, incorporating self-supervised strategies for handling unlabeled data streams. Deployment on real neuromorphic hardware, such as Intel Loihi, is also planned to assess energy efficiency and inference speed under practical conditions. Furthermore, the system may be enhanced by integrating multiple sensory modalities—such as tactile or auditory inputs—to improve environmental perception. Lastly, future work will explore combining spiking neural computation with symbolic reasoning techniques to enable robots not only to learn adaptively but also to make more informed and interpretable decisions.

Author Contributions: Zunaira Begum conceptualized the research framework, performed the literature review, developed the comparative analysis with baseline models and led the design of the proposed neuromorphic architecture. Jaibir Singh contributed to the implementation, conducted experimental evaluations, and analyzed the results across multiple datasets and also prepared the graphical and visual components of the manuscript. All authors collaborated on writing, reviewing, and editing the manuscript and approved the final version for submission.

Data availability: Data available upon request.

Conflict of Interest: There is no conflict of Interest.

Ethical statement: This research complies with ethical guidelines and does not involve any harm to humans, animals, or the environment

Funding: The research received no external funding.

Similarity checked: Yes.

References

[1] M. McCloskey and N. J. Cohen, "Catastrophic interference in connectionist networks: The sequential learning problem," Psychology of Learning and Motivation, vol. 24, pp. 109–165, 1989.

[2] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.

[3] F. Zenke, B. Poole, and S. Ganguli, "Continual learning through synaptic intelligence," in Proceedings of ICML, pp. 3987–3995, 2017.

[4] C. Liu, A. van Schaik, B. Morreale, and G. Indiveri, “Event-driven sensing for efficient perception: Combining neuromorphic vision and sparse coding,” in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada, May 2019, pp. 8681–8687, doi: 10.1109/ICRA.2019.8794402.

[5] H. Tang, Y. Li, and Y. Ma, “Neuroscience-inspired artificial intelligence,” Nature Neuroscience, vol. 24, pp. 1603–1615, 2021.

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

[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] D. Gehrig, H. Rebecq, G. Gallego, and D. Scaramuzza, “DSEC: A dataset for simultaneous stereo matching and optical flow estimation with event cameras,” in Proceedings of RSS, 2021.

[9] M. S. Lakshmi, K. S. Ramana, M. J. Pasha, K. Lakshmi, N. Parashuram, and M. Bhavsingh, “Minimizing the localization error in wireless sensor networks using multi-objective optimization techniques,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 10, no. 2s, pp. 306–312, 2022. doi: 10.17762/ijritcc.v10i2s.5948.

[10] M. Davies et al., “Loihi: A neuromorphic manycore processor with on-chip learning,” IEEE Micro, vol. 38, no. 1, pp. 82–99, 2018.

[11] M. McCloskey and N. J. Cohen, “Catastrophic interference in connectionist networks: The sequential learning problem,” Psychology of Learning and Motivation, vol. 24, pp. 109–165, 1989.

[12] J. Kirkpatrick et al., “Overcoming catastrophic forgetting in neural networks,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, Mar. 2017, doi: 10.1073/pnas.1611835114.

[13] J. Kirkpatrick et al., “Overcoming catastrophic forgetting in neural networks,” Proceedings of the National Academy of Sciences, vol. 114, no. 13, pp. 3521–3526, 2017.

[14] F. Zenke, B. Poole, and S. Ganguli, “Continual learning through synaptic intelligence,” in Proceedings of ICML, pp. 3987–3995, 2017.

[15] H. Shin et al., “Continual learning with deep generative replay,” in Proceedings of NeurIPS, pp. 2990–2999, 2017.

[16] R. Rusu et al., “Progressive neural networks,” arXiv preprint arXiv:1606.04671, 2016.

[17] P. Kumar, M. K. Gupta, C. R. S. Rao, M. Bhavsingh, and M. Srilakshmi, “A Comparative Analysis of Collaborative Filtering Similarity Measurements for Recommendation Systems,” International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 3s, pp. 184–192, Mar. 2023, doi: 10.17762/ijritcc.v11i3s.6180.

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

[19] S. B. Hassan et al., “Online learning in spiking neural networks with memory retention,” in Frontiers in Neuroscience, vol. 14, p. 217, 2020.

[20] 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.

[21] M. Orchard, C. Meyer, R. Etienne-Cummings, C. Posch, N. Thakor, and R. Benosman, “HFirst: A temporal approach to object recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 10, pp. 2028–2040, Oct. 2015.

[22] A. Amir, B. Taba, D. Berg, T. Melano, J. McKinstry, C. Di Nolfo, T. Nayak, A. Andreopoulos, G. Garreau, M. Mendoza et al., “A low power, fully event-based gesture recognition system,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7243–7252, 2017.

[23] J. Kirkpatrick et al., “Overcoming catastrophic forgetting in neural networks,” Proc. Natl. Acad. Sci. USA, vol. 114, no. 13, pp. 3521–3526, 2017.

[24] Z. Li and D. Hoiem, “Learning without forgetting,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 12, pp. 2935–2947, Dec. 2018.

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

[26] H. Shin et al., “Continual learning with deep generative replay,” in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017.