Correlation of Gradient Sparsity and Inference Latency in Adversarially Trained Multimodal SNNs on Event-Based Vision Datasets

Assignee Research

June 11, 2026

Abstract

Spiking neural networks (SNNs) have received increasing attention due to their high biological plausibility and energy efficiency. The binary spike-based information propagation enables efficient sparse computation in event-based and static computer vision applications. However, the weight precision and especially the membrane potential precision remain as high-precision values (e.g., 32 bits) in state-of-the-art SNN algorithms. Each neuron in an SNN stores the membrane potential over time and typically updates its value in every time step. Such frequent read/write operations of high-precision

1 Introduction

This paper examines: SpQuant-SNN: ultra-low precision membrane potential with sparse activations unlock the potential of on-device spiking neural networks applications. Research question: How does gradient sparsity in adversarially trained multimodal SNNs correlate with inference latency reductions on event-based vision datasets relative to standard training methods?.

2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.0/10.

3 Results

13 papers retrieved. 11 claims extracted; 11 independently verified. Quality review score: 8.0/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
Spiking neural networks (SNNs) have received increasing attention due to their high biological plausibility and energy e	✓	0.28
The binary spike-based information propagation enables efficient sparse computation in event-based and static computer v	✓	0.30
The weight precision and especially the membrane potential precision remain as high-precision values (e.g., 32 bits) in	✓	0.35
Each neuron in an SNN stores the membrane potential over time and typically updates its value in every time step.	✓	0.30
Such frequent read/write operations of high-precision membrane potential incur storage and memory access overhead in SNN	✓	0.40
Prior works have explored the time step reduction and low-precision representation of membrane potential at a limited sc	✓	0.39
Recent advances in on-device AI present pruning and quantization optimization with different architectures and datasets.	✓	0.29
Simultaneous pruning with quantization is highly under-explored in SNNs.	✓	0.24
SpQuant-SNN is a fully-quantized spiking neural network with ultra-low precision weights, membrane potential, and high s	✓	0.40
SpQuant-SNN enables the end-to-end low precision with significantly reduced operations on SNN.	✓	0.28
SpQuant-SNN proposes an integer-only quantization scheme for the membrane potential with a stack.	✓	0.23

References

- http://arxiv.org/abs/2302.00232v1
- https://arxiv.org/abs/2110.02929
- https://www.semanticscholar.org/paper/cbe9b59ab1f5c19e61540cdad4a11c8007786806