Spike-Synchrony-Dependent Learning Enhances Adversarial Robustness in Vision-Language Model Alignment

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does spike-synchrony-dependent learning affect the robustness of cross-modal alignment scores against adversarial perturbations in vision-language models evaluated on the VQA v2 dataset. 17 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Learning with Spike Synchrony in Spiking Neural Networks. Research question: How does spike-synchrony-dependent learning affect the robustness of cross-modal alignment scores against adversarial perturbations in vision-language models evaluated on the VQA v2 dataset?.

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 3.0/10.

3 Results

13 papers retrieved. 17 claims extracted; 0 independently verified. Quality review score: 3.0/10.

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
SSDP was evaluated on static image datasets including Fashion-MNIST, CIFAR-10, CIFAR-100, and ImageNet.	×	0.08
SSDP was evaluated on event-driven neuromorphic vision datasets N-MNIST and CIFAR10-DVS.	×	0.07
SSDP was evaluated on auditory high temporal datasets SHD and SSC.	×	0.07
The Proposed SpikingResformer-L achieved a Top-1 classification accuracy of 79.35$\pm$0.36 on ImageNet-1K.	×	0.03
The Proposed SpikingResformer-L has 60.38M parameters and 8.89G MACs.	×	0.01
Spikingformer-8-768 achieved a Top-1 classification accuracy of 74.81 on ImageNet-1K.	×	0.02
Spike-driven Transformer-8-768 achieved a Top-1 classification accuracy of 76.32 on ImageNet-1K.	×	0.04
DHSRNN+SSDP achieved an accuracy of 89.1$\pm$0.21 on the SHD dataset.	×	0.02
SRNN achieved an accuracy of 90.4 on the SHD dataset.	×	0.01
DHSNN+SSDP achieved an accuracy of 82.86$\pm$0.26 on the SSC dataset.	×	0.02
The Proposed SpikingResformer-Cifar achieved an accuracy of 96.24$\pm$0.29 on the CIFAR-10 dataset.	×	0.04
The Proposed SpikingResformer-Cifar achieved an accuracy of 79.48$\pm$0.27 on the CIFAR-100 dataset.	×	0.02
Transformer-4-384 achieved an accuracy of 96.73 on the CIFAR-10 dataset.	×	0.03
Transformer-4-384 achieved an accuracy of 81.02 on the CIFAR-100 dataset.	×	0.01
The synchrony gate λb,j,i equals 1 if and only if both the presynaptic neuron i and postsynaptic neuron j emit at least	×	0.04
Recording only first-spike times requires O(BC) memory.	×	0.05
Silent units are assigned a first-spike time of T, resulting in large time differences and negligible coincidence streng	×	0.03

References

- http://arxiv.org/abs/2008.07651v1
- http://arxiv.org/abs/2505.14841v2
- http://arxiv.org/abs/2306.11065v1