International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 9, Issue 9, pp 10-14, September-2025
https://doi.org/10.47001/IRJIET/2025.909002

NeuroMimicry Attacks: Adversarial Evasion in Spiking Neuromorphic Systems

[1]Alex Mathew, [2]Frank Valentin, [3]Audrey Tobesman

[1,2,3]Department of Cybersecurity, Bethany College, USA
Author's Email: amathew@bethanywv.edu

Abstract **- Neuromorphic computing has become popular in robotics, edge devices and IoT because of its energy efficiency and biological inspiration. These systems are based on spiking neural networks (SNNs), which process information in discrete spike events, providing real-time and low-power operation. However, in spite of these advantages, the safety of spiking neuromorphic systems has not been studied extensively as compared to traditional deep learning systems. In this paper, we present NeuroMimicry Attacks, a type of adversarial evasion attack in which adversarial examples are patterns of the spike-train that are highly similar to a legitimate activity but reach malicious goals. These attacks take advantage of the temporal and spatiotemporal properties of SNNs and are challenging to identify using the current anomaly detection systems. This work has four contributions: first, a taxonomy of mimicry-based adversarial attacks is created; second, algorithms to generate realistic spike-train perturbations and synthetic mimicry patterns are proposed; third, defense strategies are proposed, including spatiotemporal anomaly detection and adversarial training; fourth, the work has been experimentally validated using benchmark neuromorphic datasets and platforms. Findings indicate that the NeuroMimicry Attack is a major threat and requires strong defensive systems specific to neuromorphic systems.**

Keywords: Neuromorphic computing, spiking neural networks, adversarial attacks, mimicry, anomaly detection.

I. INTRODUCTION

Neuromorphic computing is increasingly adopted in robotics, IoT, and edge AI because it offers real-time processing with low energy consumption. Unlike conventional architectures, spiking neural networks (SNNs) process data through spike events that mimic biological neurons (Rathi *et al.,* 2023). While adversarial machine learning has received extensive attention in deep learning systems such as CNNs and transformers, limited effort addresses security in neuromorphic systems (Aitsam *et al.,* 2022). This leaves spiking architectures vulnerable to new forms of attacks. Mimicry-based adversarial strategies are particularly concerning because they operate stealthily, consume minimal energy, and adapt to dynamic environments (Kim & Kim, 2024). That said, this study raises three research questions: how can adversaries craft spike-train patterns that resemble legitimate behavior, how effective are such attacks against current neuromorphic anomaly detectors, and what defenses preserve system performance while improving resilience? The contributions are fourfold: define NeuroMimicry Attacks, propose a taxonomy and methods, design defensive strategies, and validate the findings through experiments on benchmark platforms.

II. BACKGROUND AND RELATED WORK

Spiking neural networks (SNNs) are event-driven systems where neurons communicate through discrete spikes. They encode temporal information using spike-train coding, making them suitable for real-time and low-power applications (Oh *et al.,* 2022). Research on adversarial machine learning has focused on deep learning models, showing vulnerabilities to perturbations that mislead classifiers (Rathi & Roy, 2024). Neuromorphic computing introduces unique challenges due to temporal dependencies and hardware constraints. Security research in this domain has mainly explored hardware Trojans, fault injection, and side-channel attacks, but these focus on hardware integrity rather than algorithmic vulnerabilities (Oh *et al.,* 2022; Rathi & Roy, 2024). Limited studies address stealthy attacks that manipulate spike trains while maintaining biological plausibility. Existing anomaly detection strategies are not designed to handle temporally coherent mimicry patterns (Nazari *et al.,* 2024; Oh *et al.,* 2022; Rathi & Roy, 2024). This gap highlights the need for a systematic study of adversarial evasion in spiking systems. Our work addresses this gap by defining NeuroMimicry Attacks and evaluating their impact, while also proposing defenses tailored to the dynamics of SNNs.

III. THREAT MODEL AND ATTACK SURFACE

The threat model considers adversaries with varying levels of system knowledge. In a white-box setting, the attacker has full access to synaptic weights, thresholds, and training data, enabling precise spike-train manipulation (Kudithipudi *et al.,* 2025). In a black-box model, the attacker

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 9, pp 10-14, September-2025

https://doi.org/10.47001/IRJIET/2025.909002

only observes system inputs and outputs, relying on query-based adaptation (Tong *et al.,* 2023). The gray-box case involves partial knowledge, such as network architecture, but not parameter details. Attack goals include evading anomaly detection modules, misleading decision-making tasks, and extracting sensitive information from spike patterns (Wang *et al.,* 2023). Attacks must satisfy two constraints. First, they must maintain biological plausibility so that generated spike trains resemble natural neuronal activity (Kudithipudi *et al.,* 2025; Tong *et al.,* 2023). Second, they must be energy efficient to operate within neuromorphic hardware limits. These constraints make mimicry attacks stealthier than traditional adversarial noise (Tong *et al.,* 2023). The combination of multiple adversary models, practical objectives, and strict constraints defines a broad and challenging attack surface for spiking neuromorphic systems.

IV. NEUROMIMICRY ATTACK DESIGN

NeuroMimicry Attacks are formally defined as the transformation of a legitimate spike train S into an adversarial spike train S' such that the similarity measure between S and S' exceeds a threshold τ, while S' causes misclassification or undesired system behavior (Leontev *et al.,* 2021). Several techniques support this design. Spike timing perturbation introduces minor jitter or delays that preserve biological plausibility while altering decision boundaries. Generative models such as GANs or VAEs synthesize spike-train patterns that mimic natural activity while embedding adversarial intent (Liang *et al.,* 2021). Reinforcement learning agents can iteratively adapt spike modifications to maximize evasion success (Lin *et al.,* 2022). Evaluation relies on metrics, i.e., fooling rate, temporal similarity measures like Victor–Purpura and van Rossum distance, and computational overhead. These methods provide adversaries with scalable strategies to construct mimicry attacks that remain undetected yet effective.

V. PROPOSED DEFENSES

To counter mimicry attacks, several defense strategies are proposed. Spatiotemporal anomaly detection employs graph neural networks to capture dependencies in spike patterns beyond simple temporal features (Shen *et al.,* 2021). Temporal clustering methods identify subtle irregularities in spike timing distributions. Adversarial training improves resilience by retraining models with generated mimicry samples to expose weaknesses. Bio-inspired defenses such as synaptic noise injection and structural redundancy increase robustness by adding variability that disrupts precise mimicry. Explainable methods enhance transparency by visualizing spike-train deviations that are hard to quantify with standard metrics. Combining these approaches builds a layered defense framework that balances detection accuracy with system efficiency. These defenses aim to reduce the success rate of mimicry attacks while preserving the low latency and energy efficiency that make neuromorphic computing attractive for edge applications.

VI. METHODOLOGY BLOCK DIAGRAM

The methodology is designed through a systematic review of prior literature on neuromorphic computing and adversarial machine learning. The workflow is structured into five main modules that should be included in the block diagram:

1. Literature Review: Analysis of neuromorphic systems, adversarial ML, and existing security gaps.
2. Taxonomy Development: Classification of NeuroMimicry Attacks and Adversary Models.
3. Attack Design: Formal definition and methods including spike perturbation, generative models, and reinforcement learning.
4. Defense Design: Strategies such as spatiotemporal anomaly detection, adversarial training, and bio-inspired resilience.
5. Experimental Evaluation: Testing on neuromorphic platforms with benchmark datasets.

Figure 1: Methodology block diagram

VII. METHODOLOGY BLOCK DIAGRAM

The experiments use both software simulators and neuromorphic hardware. Loihi SDK and SpiNNaker provide hardware-based evaluations, while Nengo and Brian2 simulate spike-based learning for flexibility. Three benchmark datasets are selected. NMNIST for event-based vision, DVS Gesture for temporal motion recognition, and SHD for audio spike-train data. Evaluation metrics include attack success rate, measuring how often mimicry bypasses anomaly detection, defense accuracy against mimicry, latency impact on system performance, and energy overhead introduced by defense mechanisms. This setup ensures fair and reproducible testing of both attacks and defenses across multiple domains.

VIII. RESULTS AND ANALYSIS

Results show NeuroMimicry Attacks are effective across NMNIST, DVS Gesture, and SHD datasets, just as echoed by Marchisio *et al.* (2021). Mimicry achieves high fooling rates while maintaining spike-train similarity, outperforming random adversarial noise in stealthiness. Baseline anomaly detectors struggle to identify mimicry patterns, confirming their vulnerability. Proposed defenses, especially

ISSN (online): 2581-3048
Volume 9, Issue 9, pp 10-14, September-2025
https://doi.org/10.47001/IRJIET/2025.909002

spatiotemporal detection and adversarial training, improve resilience but introduce computational and energy costs (Liu *et al.*, 2024). Trade-offs emerge between higher detection accuracy and increased latency or energy use, with hardware-specific variations. Analysis highlights that mimicry is significantly harder to detect than noise due to its temporal coherence. These findings confirm the threat potential of NeuroMimicry Attacks while providing measurable insights into defense efficiency and cost.

IX. DISCUSSION

The findings have direct implications for edge AI applications such as autonomous drones, IoT devices, and medical implants, where neuromorphic hardware may be deployed (Bharath *et al.*, 2025). Successful mimicry attacks in these domains could cause unsafe decisions or data leakage. As adoption expands, such vulnerabilities raise broader cyber-physical risks, particularly in safety-critical systems (Borra *et al.*, 2024). This work also reveals limitations. Attack performance may depend on specific neuromorphic platforms, and evaluation datasets remain limited in diversity compared to real-world conditions (Wang *et al.*, 2025). Ethical considerations are central, as adversarial research poses dual-use concerns. Responsible disclosure and defensive framing are necessary to prevent misuse. The study encourages integrating security into neuromorphic design pipelines rather than treating it as an afterthought.

X. CONCLUSION AND FUTURE WORK

In summary, this study shows that spiking neuromorphic systems are vulnerable to NeuroMimicry Attacks that evade anomaly detection through biologically plausible perturbations. Contributions include a taxonomy of mimicry attacks, formal attack design methods, defense strategies, and experimental validation across multiple datasets and platforms. Results confirm that mimicry is stealthier and more effective than random adversarial noise, demanding specialized defenses. That said, all future work should focus on hardware-level defenses that monitor energy signatures or timing irregularities, extending research to multi-modal spike-train systems, and integrating defenses with secure neuromorphic AI frameworks. Building resilience against adversarial mimicry is essential to protect edge AI deployments in robotics, IoT, and healthcare applications.

REFERENCES

[1] Aitsam, M., Davies, S., & Di Nuovo, A. (2022). Neuromorphic computing for interactive robotics: A systematic review. *IEEE Access*, 10, 122261-122279.

[2] Bharath, N., Tiwari, P., & Lakshmi, D. (2025). Sustainable AI hardware for advanced healthcare diagnostics. *In AI-Powered Systems for Healthcare Diagnostics and Treatment* (pp. 267-310). *IGI Global Scientific Publishing*.

[3] Borra, R. (2024). Neuromorphic Computing: Bridging Biological Intelligence and Artificial Intelligence. *International Journal of Engineering and Advanced Technology*, 14(2), 10-35940.

[4] Kim, E., & Kim, Y. (2024). Exploring the potential of spiking neural networks in biomedical applications: Advantages, limitations, and future perspectives. *Biomedical Engineering Letters*, 14(5), 967-980.

[5] Kudithipudi, D., Schuman, C., Vineyard, C. M., Pandit, T., Merkel, C., Kubendran, R.,... & Furber, S. (2025). Neuromorphic computing at scale. *Nature*, 637(8047), 801-812.

[6] Leontev, M., Antonov, D., & Sukhov, S. (2021, September). Robustness of spiking neural networks against adversarial attacks. *In the 2021 International Conference on Information Technology and Nanotechnology (ITNT)* (pp. 1-6). IEEE.

[7] Liang, L., Hu, X., Deng, L., Wu, Y., Li, G., Ding, Y.,... & Xie, Y. (2021). Exploring adversarial attack in spiking neural networks with spike-compatible gradient. *IEEE Transactions on Neural Networks and Learning Systems*, 34(5), 2569-2583.

[8] Lin, X., Dong, C., Liu, X., & Cheng, D. (2022, December). Spiking neural networks subject to adversarial attacks in the spiking domain. *In International Conference on Machine Learning for Cyber Security* (pp. 457-471). Cham: Springer Nature Switzerland.

[9] Liu, R., Shi, J., Chen, X., & Lu, C. (2024). Network anomaly detection and security defense technology based on machine learning: A review. *Computers and Electrical Engineering*, 119, 109581.

[10] Marchisio, A., Pira, G., Martina, M., Masera, G., & Shafique, M. (2021, July). Dvs-attacks: Adversarial attacks on dynamic vision sensors for spiking neural networks. *In 2021 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-9). IEEE.

[11] Nazari, N., Gubbi, K. I., Latibari, B. S., Chowdhury, M. A., Fang, C., Sasan, A., ... & Salehi, S. (2024, May). Securing on-chip learning: Navigating vulnerabilities and potential safeguards in spiking neural network architectures. *In 2024 IEEE International Symposium on Circuits and Systems (ISCAS)* (pp. 1-5). IEEE.

[12] Oh, S., Kwon, D., Yeom, G., Kang, W. M., Lee, S., Woo, S. Y., ... & Lee, J. H. (2022). Neuron circuits for low-power spiking neural networks using time-to-first-spike encoding. *IEEE Access*, 10, 24444-24455.

ISSN (online): 2581-3048
Volume 9, Issue 9, pp 10-14, September-2025
https://doi.org/10.47001/IRJIET/2025.909002

[13] Rathi, N., Chakraborty, I., Kosta, A., Sengupta, A., Ankit, A., Panda, P., & Roy, K. (2023). Exploring neuromorphic computing based on spiking neural networks: Algorithms to hardware. *ACM Computing Surveys*, 55(12), 1-49.

[14] Rathi, N., & Roy, K. (2024). Lite-snn: Leveraging inherent dynamics to train energy-efficient spiking neural networks for sequential learning. *IEEE Transactions on Cognitive and Developmental Systems*, 16(6), 1905-1914.

[15] Tong, C., Zheng, X., Li, J., Ma, X., Gao, L., & Xiang, Y. (2023). Query-efficient black-box adversarial attacks on automatic speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31, 3981-3992.

[16] Shen, J., Liu, J. K., & Wang, Y. (2021). Dynamic spatiotemporal pattern recognition with recurrent spiking neural network. *Neural Computation*, 33(11), 2971-2995.

[17] Wang, H., Li, Y. F., & Gryllias, K. (2023). Brain-inspired spiking neural networks for industrial fault diagnosis: A survey, challenges, and opportunities. *arXiv preprint arXiv:*2401.02429. https://doi.org/10.48550/arXiv.2401.02429.

[18] Wang, T., Guo, J., Zhang, B., Yang, G., & Li, D. (2025). Deploying AI on edge: Advancement and challenges in edge intelligence. *Mathematics*, 13(11), 1878.

AUTHORS BIOGRAPHY

Alex Mathew, Ph.D., CISA, CISSP, MCSA, CEH, CHFI, ECSA, CEI,CCNP

Is an Associate Professor in the Department of Cybersecurity at Bethany College (West Virginia, USA) and is widely recognized for his deep expertise in cybersecurity, cybercrime investigations, next-generation networks, data science, and IoT Azure solutions. His proficiency in security best practices, particularly in IoT, cloud systems, and healthcare IoT, is complemented by his comprehensive knowledge of industry standards such as ISO 17799, ISO 31000, ISO/IEC 27001/2, and HIPAA regulations. His credentials, including certifications in Cybersecurity and Data Science from Harvard University, further strengthen his expertise in the field.

As a certified Information systems security professional (CISSP), Mathew's leadership is evident in his role as a consultant across international regions, including India, Asia, Cyprus, and the Middle East. His extensive two-decade career, distinguished by numerous certifications and over 100 scholarly publications, underscores his commitment to advancing the field. Mathew has been a pivotal force in organizing cybersecurity conferences and establishing incubation centers, contributing significantly to the academic and professional community.

A highly sought-after speaker, Mathew's influence extends to international conferences where he shares his insights on cybersecurity, technology, and data science. His remarkable interpersonal skills and openness enhanced his ability to engage and inspire diverse audiences, further cementing his position as a leader in his field.

Audrey Tobesman, B.S. Student

Audrey Tobesman is an undergraduate student at Bethany College (West Virginia, USA), majoring in computer science with a strong interest in cybersecurity, artificial intelligence, and computer networks. Throughout her studies, she has gained experience in programming, data structures, and network security, which has fueled her curiosity about the vulnerabilities of emerging technologies.

Her research, NeuroMimicry Attacks: Adversarial Evasion in Spiking Neuromorphic Systems, explores the intersection of neuroscience-inspired computing and adversarial machine learning, a cutting-edge area that combines both theoretical and applied computer science. By analyzing how adversarial evasion techniques can exploit the unique properties of spiking neuromorphic systems, she seeks to highlight both the potential risks and the need for stronger defenses in next-generation AI architectures.

Tobesman's work reflects her broader academic goal of contributing to the advancement of secure and trustworthy artificial intelligence systems, while also preparing for a future career in cybersecurity and technology research.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 9, pp 10-14, September-2025

https://doi.org/10.47001/IRJIET/2025.909002

Valentin Aguirre Frank, BA. Student

Valentin Aguirre Frank is an undergraduate studying at Bethany College (West Virginia, USA) from Buenos Aires, Argentina. Majoring in Computer Science with interests in Cyber Security, Computer Networks and Hardware. Throughout his studies and working experience he gained knowledge in Programming, Networks, Data Structures, Cyber-attacks and Security. His experience gave him a better understanding of ways to identify Cyber-attacks and to keep systems secure from the outside of an organization.

Citation of this Article:

Alex Mathew, Frank Valentin, & Audrey Tobesman. (2025). NeuroMimicry Attacks: Adversarial Evasion in Spiking Neuromorphic Systems. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(9), 10-14. Article DOI https://doi.org/10.47001/IRJIET/2025.909002

****

