

A Unified Neural-Symbolic Intelligence Model for Real-Time Threat Detection, Incident Response Orchestration, and Critical Infrastructure Defense

Authors

Dada Jay, Godwin Frey, Asher Noah, Abi Cit, Abbas Ahsun

Date; October 6, 2026

Abstract

Critical infrastructure sectors face increasingly sophisticated cyber threats that bypass conventional signature-based defenses, yet existing AI-driven security solutions remain fragmented across detection, response, and orchestration functions. This study addresses the critical gap in unified architectures capable of simultaneously delivering real-time threat detection, automated incident response, and explainable decision-making for operational technology environments. We propose a neural-symbolic intelligence model that integrates graph neural networks, transformer-based temporal reasoning, and ontology-driven symbolic inference within a modular agentic architecture. The framework was validated using a retrospective analysis of the CICIDS2017, UNSW-NB15, and CICIOT2023 benchmark datasets, supplemented by simulated multi-stage attack scenarios representative of critical infrastructure threats. The unified model achieved 89.4% detection accuracy for coordinated multi-vector attacks, outperforming baseline deep learning models by 12.7 percentage points while reducing mean time to response by 43% compared to rule-based security orchestration systems. Feature importance analysis identified temporal dependency patterns and cross-domain correlation signals as the most influential

predictors. The study contributes a replicable architecture for integrating symbolic reasoning with neural perception in security operations, offering practitioners a pathway toward transparent, autonomous, and context-aware critical infrastructure defense. Implications extend to standards development, SOC workflow redesign, and future research on adversarial robustness in hybrid AI systems.

Keywords: Neural-symbolic AI, Threat detection, Incident response orchestration, Critical infrastructure security

1. Introduction

1.1 Background

The digital transformation of critical infrastructure sectors—energy, water, transportation, and healthcare—has created unprecedented attack surfaces that adversaries increasingly exploit through sophisticated, multi-stage campaigns . Traditional security architectures built on signature-based detection and manual incident response have proven inadequate against zero-day exploits, polymorphic malware, and advanced persistent threats that evolve faster than static defensive rules can adapt . The average time to detect and contain a breach now exceeds 200 days, a metric that underscores the fundamental limitations of reactive security postures .

In response, the security industry has witnessed a proliferation of AI-driven solutions. Deep learning models, particularly those employing transformer architectures and graph neural networks, have demonstrated impressive detection capabilities on benchmark datasets, with reported accuracies exceeding 96% on curated intrusion detection corpora . Concurrently, Security Orchestration, Automation, and Response (SOAR) platforms have emerged to streamline incident response workflows, reducing manual triage burdens through playbook automation . However, these advances have proceeded largely in isolation, producing a fragmented ecosystem where detection, response orchestration, and knowledge representation operate as disconnected components.

1.2 Problem Statement

The integration of neural learning with symbolic reasoning has emerged as a promising research direction, with recent surveys documenting significant advances in neuro-symbolic architectures across domains including robotics, healthcare, and autonomous systems . Yet the application of neural-symbolic methods to cybersecurity operations remains nascent and under-theorized. Existing AI-based intrusion detection systems, while achieving high benchmark accuracy, function as opaque "black boxes" that provide predictions without justifications—a critical deficiency in security operations centers where analysts must understand *why* an alert was triggered before authorizing containment actions . Conversely, SOAR platforms, even when augmented with

machine learning components, lack the structured knowledge representation necessary for adaptive, context-aware response selection across heterogeneous IT/OT environments .

Three interconnected limitations characterize the current state of the art. First, no unified architecture exists that seamlessly integrates neural perception (threat detection from raw telemetry) with symbolic reasoning (ontology-driven response selection and explainable decision-making). Second, existing solutions evaluate detection and response as separate problems, failing to account for the end-to-end latency and accuracy requirements of real-time critical infrastructure defense. Third, the explainability deficit in deep learning-based security tools creates adoption barriers in high-stakes operational environments where accountability and auditability are non-negotiable requirements .

1.3 Objectives of the Study

General objective: To design and validate a unified neural-symbolic intelligence model for real-time threat detection, incident response orchestration, and critical infrastructure defense.

Specific objectives:

1. To identify the architectural components and integration patterns necessary for combining neural perception with symbolic reasoning in security operations contexts.
2. To design a hybrid model that achieves superior detection accuracy for coordinated multi-vector attacks while maintaining sub-second inference latency.
3. To validate the framework using benchmark intrusion detection datasets and simulated critical infrastructure attack scenarios, comparing performance against established baselines.

1.4 Research Questions

1. What combination of neural architectures and symbolic knowledge representations most effectively supports unified detection-response workflows in critical infrastructure environments?
2. How does the proposed neural-symbolic framework compare to traditional deep learning and rule-based SOAR approaches in terms of detection accuracy, false positive reduction, and response latency?
3. What implementation barriers and design trade-offs characterize the deployment of neural-symbolic security systems in resource-constrained operational technology contexts?

1.5 Significance of the Study

For security operations practitioners and critical infrastructure administrators, this research provides a concrete architectural blueprint for transitioning from fragmented, alert-driven security postures to unified, intelligence-driven defense systems. The framework addresses the alert fatigue

problem that plagues modern SOC's by reducing false positives and automating context-aware response selection .

For policymakers and standards bodies, the study contributes evidence supporting the development of guidelines for AI deployment in critical infrastructure, particularly regarding explainability requirements and human oversight mechanisms. The findings inform ongoing discussions around NIS2 implementation and IEC 62443 compliance .

For academic literature, the research extends neural-symbolic AI theory into the cybersecurity domain, demonstrating how symbolic knowledge representation can enhance both the accuracy and interpretability of neural security models. The framework contributes to the growing body of work on hybrid AI architectures while addressing domain-specific challenges of adversarial environments and real-time constraints.

For future researchers, the study identifies open problems in adversarial robustness, cross-domain generalization, and scalable knowledge graph construction for security ontologies, providing a foundation for subsequent investigations.

1.6 Scope and Limitations

This study focuses on network-level threat detection and incident response orchestration for critical infrastructure environments, with emphasis on IT/OT convergence scenarios. The geographic scope is global in terms of dataset and standards applicability, though validation relies primarily on publicly available benchmark datasets that may not fully represent the operational characteristics of specific critical infrastructure sectors.

The study excludes endpoint-level detection, physical security integration, and human factors analysis beyond the explainability requirements for analyst trust. The use of simulated attack scenarios, while necessary for controlled evaluation of multi-stage threats, introduces limitations regarding the fidelity of adversarial behavior representation. The framework assumes availability of structured threat intelligence and security ontologies; organizations lacking such foundational resources may require additional preparatory work before implementation.

2. Literature Review

2.1 Conceptual Review

Neural-Symbolic AI refers to architectures that integrate sub-symbolic learning (neural networks) with symbolic reasoning (logic, ontologies, knowledge graphs) to combine the pattern recognition strengths of deep learning with the interpretability and compositional reasoning capabilities of symbolic systems . In security contexts, this integration enables models that both detect threats from high-dimensional telemetry and explain their decisions through structured, human-interpretable representations.

Real-Time Threat Detection encompasses the continuous monitoring, analysis, and identification of malicious activities within network and system telemetry, with latency requirements measured in milliseconds to seconds for effective intervention . The construct includes detection accuracy, false positive rate, and temporal sensitivity as key dimensions.

Incident Response Orchestration denotes the automated coordination of security actions— isolation, containment, remediation, and recovery—based on detected threats and organizational policies. In SOAR systems, orchestration involves playbook execution, tool integration, and workflow management .

Critical Infrastructure Defense refers to the protection of systems and assets so vital that their incapacitation would have debilitating effects on national security, economic stability, or public safety. The construct encompasses both technical controls and governance frameworks, including standards such as NIST CSF and IEC 62443 .

2.2 Theoretical Framework

The proposed model draws upon three theoretical foundations. **Dual-Process Theory**, as articulated by Kahneman, distinguishes between fast, intuitive (System 1) and slow, deliberative (System 2) cognitive processes . In the neural-symbolic security context, neural networks embody System 1 capabilities—rapid pattern recognition from raw telemetry—while symbolic reasoning represents System 2—deliberative, rule-governed analysis for response selection and explanation.

Autonomic Computing Theory, with its emphasis on self-management properties (self-configuration, self-healing, self-optimization, self-protection), provides a framework for understanding how SOAR systems can achieve varying degrees of autonomy . The neural-symbolic model operationalizes autonomic principles by delegating detection and initial triage to neural components while reserving response authorization and policy enforcement to symbolic reasoning.

Ontology-Driven Decision Support theory holds that structured knowledge representations enable more consistent and auditable decisions than purely procedural approaches . The use of cybersecurity ontologies such as MITRE D3FEND provides the semantic foundation for mapping detected threats to appropriate countermeasures, ensuring that response selection follows established defensive frameworks rather than ad hoc heuristics.

2.3 Empirical Review

Recent empirical work has advanced both neural detection and SOAR automation, though integration remains limited. Bhuiyan et al. (2026) proposed CrisisAI-MIS, a unified multi-modal intelligence framework for real-time crisis prediction and management, demonstrating the feasibility of integrating heterogeneous data sources for emergency response scenarios . Their architecture, however, focused on crisis management rather than cybersecurity, and the symbolic reasoning component remained underdeveloped.

CyberGuardians, introduced by researchers in 2025, achieved 96.3% accuracy and 95.7% F1 score on UNSW-NB15 and CICIDS2017 using a combination of transformer-based temporal reasoning, graph neural networks, and contrastive self-supervision . While demonstrating the power of multi-modal neural architectures, the system provided no mechanisms for explainable decision-making or response orchestration.

An AI-driven SOAR framework integrating large language models with intelligent agents reported enhanced detection accuracy and reduced human intervention on CICIDS-2017 . The LLM-based reasoning component offered a form of symbolic-like contextualization, but lacked formal ontological grounding, creating risks of hallucination and inconsistent policy enforcement.

CyberGuardX, employing a hybrid GNN-Transformer architecture with Graph Attention Anomaly Score, achieved 99.2% F1 on rare threat classes and 87% false positive reduction compared to commercial systems . The system's SHAP-based explainability provided feature attribution, but this post-hoc explanation differs fundamentally from the *a priori* symbolic reasoning that enables compositional generalization to novel threat scenarios.

Huang et al. (2025) proposed an ontology-driven approach to SOAR using MITRE-D3FEND, demonstrating how intent-based networking principles could formalize response selection . Their work established the theoretical foundation for symbolic response orchestration but did not integrate neural detection components or evaluate end-to-end system performance.

2.4 Research Gap

No validated unified neural-symbolic framework exists that integrates real-time threat detection, explainable decision-making, and automated incident response orchestration for critical infrastructure environments. Existing work has either advanced neural detection without symbolic reasoning capabilities , developed SOAR automation without robust neural perception , or theorized ontology-driven response without empirical validation in operational contexts . The present study fills this gap by designing, implementing, and validating an integrated architecture that addresses the full detection-to-response pipeline while maintaining the explainability and auditability required for critical infrastructure deployment.

3. Methodology

3.1 Research Design

This study employs a **design-based research** methodology combined with **retrospective quantitative analysis**. Design-based research is appropriate for investigating complex, real-world problems where the goal is to develop and refine an intervention (the neural-symbolic model) within its natural context . The retrospective component enables rigorous evaluation against established benchmark datasets while the design component facilitates iterative refinement of the architecture based on performance observations.

3.2 Study Area / Population

The study targets network traffic and system telemetry from critical infrastructure environments, with particular emphasis on IT/OT convergence scenarios. The population comprises network flows, system logs, and threat intelligence indicators representative of attacks against industrial control systems, energy distribution networks, and water treatment facilities. Benchmark datasets were selected for their established use in intrusion detection research and their coverage of attack categories relevant to critical infrastructure threats.

3.3 Sample Size and Sampling Technique

The study utilized three benchmark datasets: CICIDS2017 (approximately 2.8 million records across 14 attack categories), UNSW-NB15 (approximately 2.5 million records across 9 attack categories), and CICIoT2023 (approximately 45 million records across 33 attack types). Stratified random sampling was employed to ensure balanced representation of attack categories, with 70% of data allocated to training, 15% to validation, and 15% to held-out testing. The stratification ensured that rare attack classes—particularly those representative of multi-stage APT campaigns—received adequate representation during model development.

3.4 Data Collection Methods

Data were sourced from publicly available benchmark repositories maintained by the Canadian Institute for Cybersecurity and the Australian Centre for Cyber Security. These datasets provide labeled network traffic captures, flow-level features, and attack annotations. Additionally, simulated multi-stage attack scenarios were constructed using the MITRE ATT&CK framework to represent coordinated campaigns that span multiple detection categories. Simulation was necessary because existing benchmarks, while comprehensive for single-stage attacks, inadequately represent the temporal dependencies and cross-domain correlations characteristic of advanced persistent threats targeting critical infrastructure.

3.5 Research Instruments

The implementation employed Python 3.11 with PyTorch 2.1 for neural components, the DomiKnowS framework for neuro-symbolic integration, and custom ontology modules based on MITRE D3FEND and ATT&CK taxonomies. Preprocessing included standard scaling for numerical features, one-hot encoding for categorical variables, and temporal sequence construction with sliding windows of 10 time steps for transformer input. Graph construction for the GNN component followed the approach described in CyberGuardX, with inter-host relationships derived from communication patterns and shared resource access.

3.6 Validity and Reliability

Content validity was established through alignment of the threat taxonomy with MITRE ATT&CK and D3FEND frameworks, ensuring coverage of attack techniques and defensive countermeasures relevant to critical infrastructure.

Predictive validity was assessed through cross-dataset evaluation, testing models trained on CICIDS2017 against UNSW-NB15 and CICIoT2023 to evaluate generalization beyond the training distribution.

Inter-rater reliability for attack scenario labeling was ensured through independent classification by three security analysts with CISSP certifications, achieving Cohen's κ of 0.87.

3.7 Data Analysis Techniques

The neural-symbolic model was compared against four baselines: (1) a transformer-only detection model, (2) a GNN-only detection model, (3) a rule-based SOAR system without neural components, and (4) a commercial IDS baseline represented by published performance metrics. Performance metrics included accuracy, precision, recall, F1 score, false positive rate, and mean time to response (MTTR). Five-fold cross-validation was employed for hyperparameter tuning, with the held-out test set reserved for final evaluation. Statistical significance of performance differences was assessed using paired t-tests with Bonferroni correction. Feature importance for neural components was computed using SHAP values, while symbolic reasoning traces were analyzed for consistency with D3FEND-recommended countermeasures.

3.8 Ethical Considerations

This study utilized only de-identified, publicly available benchmark datasets. No protected health information, personally identifiable information, or proprietary industrial data were accessed. The research protocol was submitted to the institutional review board and determined to qualify for exemption under 45 CFR 46.104(d)(4) as research involving only publicly available data. All simulated attack scenarios were conducted in isolated laboratory environments with no connection to production systems.

4. Results

4.1 Data Presentation

Table 1 presents descriptive statistics for the benchmark datasets employed in this study.

Table 1. Benchmark Dataset Characteristics

Dataset	Records (millions)	Attack Categories	Benign:Attack Ratio	Temporal Span
CICIDS2017	2.8	14	4.2:1	5 days

Dataset	Records (millions)	Attack Categories	Benign:Attack Ratio	Temporal Span
UNSW-NB15	2.5	9	3.8:1	2 days
CICIoT2023	45.0	33	6.1:1	14 days

Table 2 reports the class distribution for attack categories most relevant to critical infrastructure threats, aggregated across datasets.

Table 2. Critical Infrastructure-Relevant Attack Distribution

Attack Category	Count	Percentage
Reconnaissance	1,247,892	24.3%
Initial Access	892,441	17.4%
Lateral Movement	654,218	12.7%
Command & Control	521,336	10.1%
Impact/Disruption	418,775	8.2%
Other Categories	1,398,441	27.3%

4.2 Analysis of Results

The unified neural-symbolic model achieved **89.4% detection accuracy** for coordinated multi-vector attacks across the combined test sets, representing a **12.7 percentage point improvement** over the best-performing baseline (transformer-only model, 76.7% accuracy). Table 3 presents the comparative performance of all evaluated approaches.

Table 3. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score	FPR	MTTR (s)
Neural-Symbolic (Proposed)	89.4%	91.2%	87.8%	89.5%	3.2%	4.7
Transformer-only	76.7%	78.4%	74.9%	76.6%	8.9%	12.3
GNN-only	74.2%	75.8%	72.1%	73.9%	10.4%	11.8
Rule-based SOAR	61.3%	58.7%	63.2%	60.9%	15.7%	8.2

The improvement over the transformer-only baseline was statistically significant ($t = 8.42$, $p < .001$, Cohen's $d = 1.37$), indicating a large effect size. False positive rate was reduced by 64% compared to the best neural baseline (3.2% vs. 8.9%), directly addressing the alert fatigue problem in security operations.

Feature importance analysis using SHAP values identified the top predictors as: temporal dependency patterns across event sequences (mean $|\text{SHAP}| = 0.184$), cross-domain correlation signals between network and system telemetry (0.156), and graph structural features representing inter-host communication anomalies (0.143). Symbolic reasoning traces demonstrated 94.3% consistency with D3FEND-recommended countermeasures for detected threat categories, with deviations occurring primarily in novel attack scenarios not represented in the ontology.

5. Discussion

5.1 Interpretation

The finding that neural-symbolic integration substantially outperforms neural-only approaches for coordinated multi-vector attacks (89.4% vs. 76.7%) answers Research Question 2 and aligns with theoretical expectations regarding the complementary strengths of neural and symbolic components. Neural networks excel at detecting low-level anomalies in high-dimensional telemetry, but struggle to maintain consistent decision boundaries across temporally extended attack campaigns where individual events may appear benign in isolation. Symbolic reasoning, grounded in threat ontologies, provides the compositional structure necessary for correlating disparate signals into coherent threat assessments.

The 64% reduction in false positive rate compared to neural baselines addresses a critical operational barrier to AI adoption in SOC environments. Alert fatigue from excessive false positives is consistently identified as a primary cause of analyst burnout and missed detections . The neural-symbolic architecture mitigates this by requiring both neural anomaly detection *and* symbolic consistency with known attack patterns before generating high-confidence alerts, effectively using the ontology as a filtering mechanism.

The MTTR reduction from 12.3 seconds (transformer baseline) to 4.7 seconds (proposed model) validates the response orchestration component. This improvement stems from the symbolic reasoning module's ability to select appropriate countermeasures directly from the D3FEND ontology rather than requiring separate classification or playbook matching steps. The integration of the CrisisAI-MIS framework informed this design, particularly its approach to unified multi-modal intelligence for real-time response coordination.

5.2 Implications

Academic implications: This study extends neural-symbolic AI theory into the cybersecurity domain, demonstrating that the integration benefits documented in robotics and healthcare transfer to adversarial environments with real-time constraints. The finding that symbolic consistency checking reduces false positives while neural perception maintains detection sensitivity establishes a design pattern applicable beyond security—any domain requiring both pattern recognition and rule-governed decision-making can potentially benefit from this architecture.

Practical implications: Critical infrastructure administrators should prioritize investment in ontology development and knowledge graph construction as prerequisites for neural-symbolic security deployment. The 89.4% accuracy figure, while not matching the highest benchmark scores reported for pure neural models on single-stage attacks , represents superior performance for the multi-stage, coordinated threats that actually threaten critical infrastructure. Organizations should monitor mean time to detect (MTTD), MTTR, and false positive rate as primary performance indicators, targeting MTTR below 10 seconds for automated response and below 60 seconds for human-in-the-loop workflows.

5.3 Limitations

1. **Benchmark representativeness:** The benchmark datasets, while standard in the field, may not fully capture the operational characteristics of specific critical infrastructure sectors such as energy SCADA systems or water treatment networks. Generalization to these contexts requires domain-specific validation.
2. **Simulated multi-stage attacks:** The simulated attack scenarios, though informed by MITRE ATT&CK, represent constructed rather than observed adversarial behavior. Real-world adversaries may employ techniques and sequences not represented in the simulation.

3. **Ontology coverage:** The symbolic reasoning component depends on the completeness of the D3FEND and ATT&CK ontologies. Novel attack techniques absent from these frameworks will not be matched to countermeasures, potentially limiting response effectiveness.
4. **Computational requirements:** The neural-symbolic architecture demands greater computational resources than lightweight IDS solutions, potentially limiting deployment in resource-constrained OT environments without edge acceleration.

5.4 Future Research Directions

1. **Sector-specific validation:** Extension of the framework to energy, water, and healthcare sectors with domain-specific ontologies and telemetry characteristics.
2. **Adversarial robustness testing:** Systematic evaluation of neural-symbolic models against adversarial examples designed to exploit the neural perception component while evading symbolic consistency checks.
3. **Continuous ontology learning:** Development of methods for automatically extending security ontologies based on novel threat observations, reducing dependence on manual knowledge engineering.
4. **Human factors integration:** Investigation of how explainability mechanisms affect analyst decision-making quality and speed in operational SOC environments.

6. Conclusion

This study designed and validated a unified neural-symbolic intelligence model achieving 89.4% detection accuracy for coordinated multi-vector attacks while reducing false positives by 64% and mean time to response by 62% compared to neural-only baselines. The main contribution is a replicable architecture that integrates graph neural networks, transformer temporal reasoning, and ontology-driven symbolic inference within a modular agentic framework, demonstrating that symbolic knowledge representation can enhance both the accuracy and interpretability of AI-driven security operations. For critical infrastructure administrators, the practical takeaway is that investment in security ontologies and knowledge graphs yields measurable operational benefits when paired with advanced neural detection components, enabling a transition from alert-driven to intelligence-driven defense postures. As adversarial capabilities continue to evolve and critical infrastructure digitization accelerates, unified neural-symbolic architectures offer a principled pathway toward autonomous, transparent, and context-aware defense systems that can operate at machine speed while maintaining the accountability required for high-stakes operational environments.

References

1. Bhuiyan, M. S., Hasan, R., Ahmed, M. J., Hasan, M., Ahmed, K. R., & Islam, R. (2026, April). CrisisAI-MIS: A unified multi-modal intelligence framework for real-time crisis prediction and management. In *2026 IEEE 2nd International Conference on Quantum Photonics, Artificial Intelligence & Networking (QPAIN)* (pp. 1-6). IEEE.
2. Colelough, B. C., & Regli, W. (2024). Neuro-symbolic AI in 2024: A systematic review. In *1st International Workshop on Logical Foundations of Neuro-Symbolic AI (LNSAI 2024)*. <https://arxiv.org/abs/2501.05435>
3. Huang, Z., Robin, J., Herbaut, N., Ben Rabah, N., & Le Grand, B. (2025). Toward an intent-based and ontology-driven autonomic security response in Security Orchestration Automation and Response. In *Proceedings of the 20th International Conference on Availability, Reliability and Security* (pp. 1-15). Association for Computing Machinery. https://doi.org/10.1007/978-3-032-15140-7_15
4. Kordjamshidi, P. (2026). Complex reasoning over vision and language: Leveraging neurosymbolic AI. *Proceedings of the AAAI Conference on Artificial Intelligence*, *40*(46), 41326. <https://doi.org/10.1609/aaai.v40i46.41326>
5. Kulesza, K., & Zgud, A. (2024). Artificial intelligence: Concepts and interpretation. *Formy*, (22).
6. Manocchio, L. D., Layeghy, S., Lo, W. W., & Portmann, M. (2024). Flowtransformer: A transformer framework for flow-based network intrusion detection systems. *Expert Systems with Applications*, *241*, 122564.
7. Martínez, R., et al. (2026). POSEIDON: A human-centered resilience framework for critical IT/operational technology infrastructures—A case study on a high-demand operational asset. *IEEE Access*, *14*, 11578157.
8. Peratikou, A., Charalambous, E., Smyrli, P., & Stavrou, S. (2026). CY-TRUST: A sectorial SOC framework for enhanced national and cross-border cybersecurity resilience. In *Critical Information Infrastructures Security* (pp. 1-18). Springer. https://doi.org/10.1007/978-3-032-19540-1_19
9. Sikos, L. F. (2023). Cybersecurity knowledge graphs. *Knowledge and Information Systems*, *65*(9), 3511-3531.
10. Strom, B. E., et al. (2017). Finding cyber threats with ATT&CK-based analytics. *The MITRE Corporation*, Technical Report No. MTR170202.
11. Vyas, S., Hannay, J., Bolton, A., & Burnap, P. P. (2023). Automated cyber defence: A review. *arXiv preprint arXiv:2303.04926*.
12. Wang, M., Yang, N., & Weng, N. (2023). Securing a smart home with a transformer-based IoT intrusion detection system. *Electronics*, *12*(9), 2100.

13. Xie, J., Chen, Z., Zhang, R., Wan, X., & Li, G. (2024). Large multimodal agents: A survey. *arXiv preprint arXiv:2402.15116*.
14. Zhang, Y., et al. (2026). CyberGuardX: AI-optimized real-time threat detection and alert system. In *2026 IEEE International Conference on Recent Trends in Engineering, Technology and Management (ICRTETM)* (pp. 1-6). IEEE.
15. Zhao, Y., et al. (2025). Toward transparent cyber threat detection: A systematic literature review on the role of explainable AI (XAI) in information security risk management (2018-2025). In *2025 IEEE International Conference on Computer Science and Information Technology (ICCSIT)* (pp. 1-8). IEEE.