iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.84001

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Survey on Artificial Intelligence for Cyber Threat Detection and Response in Cloud Environments

Mohammed Adeen Khurshid[1], Mohammed Zuber Mulla[2], Mohammed Aziz A Khazi[3], Moomin Mukhtar[4]
HKBK College of Engineering, Bengaluru, Karnataka

Abstract: ***The rapid expansion of cloud computing infrastructures has fundamentally transformed how organizations manage and deploy digital services, simultaneously introducing a complex and evolving attack surface that traditional security mechanisms fail to adequately address. This survey examines the convergence of artificial intelligence (AI) and autonomous cybersecurity with a focus on cloud environments. We systematically review thirteen recent papers spanning five core research themes: AI-driven threat detection and classification, explainable AI (XAI) for cybersecurity transparency, autonomous response and mitigation strategies, real-time cyber threat attribution, and AI-enhanced education for cybersecurity workforce development. Our analysis highlights the state-of-the-art techniques including Graph Neural Networks (GNNs), transformer-based attention mechanisms, Federated Deep Learning (FDL), reinforcement learning, and multi-modal data fusion, all applied to the challenge of building self-healing, autonomous cloud defense systems. We further discuss persistent challenges such as dataset quality, model interpretability, adversarial robustness, and the gap between academic research and real-world deployment. This survey provides a structured synthesis of the current state of the art, identifying key research directions for the next generation of intelligent, autonomous cloud security systems.***

Keywords: Artificial intelligence, cloud security, autonomous cyber defense, threat detection, explainable AI, reinforcement learning, federated learning, intrusion detection, deep learning, anomaly detection.

I. INTRODUCTION

Cloud computing has emerged as the dominant paradigm for modern digital infrastructure, offering scalability, flexibility, and cost efficiency across industries (akbari2024cloud?; afraji2025ddos?). However, this rapid adoption has introduced an expanded and increasingly sophisticated attack surface. Cyber threats including Distributed Denial-of-Service (DDoS) attacks, ransomware, advanced persistent threats (APTs), insider threats, and zero-day exploits have grown in both frequency and complexity, exposing critical services, sensitive data, and operational systems to significant risk (swetha2025ai?; sharma2024cloud?).

Traditional security mechanisms — signature-based intrusion detection systems, rule-based firewalls, and static perimeter defenses — are fundamentally reactive and cannot keep pace with novel, polymorphic, or AI-augmented attack strategies (mishra2025ai?; krishnappa2024cloud?). The limitations of human-driven security operations, compounded by a global cybersecurity talent shortage, underscore the urgent need for intelligent, autonomous defense systems that can detect, classify, and respond to threats without constant human supervision (tolah2026teachsecure?).

AI has emerged as a transformative force in cybersecurity, enabling systems to learn from data, adapt to new attack patterns, and make autonomous decisions at machine speed (swetha2025ai?; sharma2025xai?). Within cloud environments specifically, AI-driven approaches offer the ability to process high-volume telemetry streams, identify anomalies across heterogeneous workloads, and enforce dynamic security policies in real time (nutalapati2025cloud?; akbari2024cloud?).

This survey reviews thirteen recent papers spanning the period 2024–2026 to synthesize the state of the art in AI-driven cybersecurity for cloud environments. Our review is structured around five thematic areas:

1) AI-based threat detection and classification in cloud environments
2) Explainable AI (XAI) for cybersecurity transparency and trust
3) Autonomous threat response and mitigation strategies
4) Real-time cyber threat attribution using AI
5) AI-enhanced cybersecurity education and workforce development

The remainder of this paper is organized as follows. Section 2 provides background on cloud security challenges. Sections 3–7 cover the five thematic review areas. Section 8 presents a comparative analysis of surveyed works. Section 9 discusses open challenges and future directions. Section 10 concludes the paper.

II. BACKGROUND: CLOUD SECURITY CHALLENGES

A. Cloud Computing Models and Security Implications

Cloud computing is delivered across three primary service models — Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) — and four deployment models: public, private, community, and hybrid clouds. Each model carries distinct security implications (akbari2024cloud?). Hybrid cloud configurations, while offering operational flexibility, introduce data fragmentation, inconsistent security policies across providers, and compliance complexity that attackers can exploit (srimathi2023multicloud?).

B. The Evolving Threat Landscape

The threat landscape facing cloud environments is both broad and rapidly evolving. Key threat categories include:

DDoS Attacks: Volumetric, protocol-based, and application-layer attacks that exploit the cloud's shared infrastructure to amplify disruption (afraji2025ddos?). The use of botnets composed of compromised IoT devices has intensified this threat.

Ransomware and Malware: Modern ransomware employs polymorphic encryption, living-off-the-land tactics, and AI-generated payloads that evade signature-based detection (prajwala2025ransomware?). Edge and cloud systems face particular vulnerability due to their distributed execution environments.

Advanced Persistent Threats (APTs): Long-duration, stealthy intrusions that leverage lateral movement across virtual machines and containers (krishnappa2024cloud?). Multi-stage APTs exploit the multi-tenancy and shared resource model intrinsic to cloud platforms.

Zero-Day Exploits: Previously unknown vulnerabilities for which no signature or patch exists, requiring behavioral rather than pattern-based detection (nutalapati2025cloud?).

Insider Threats and Privilege Escalation: Authorized users abusing legitimate access, or attackers escalating privileges after initial compromise (mishra2025ai?).

C. Limitations of Traditional Defenses

Signature-based IDS cannot detect zero-day or polymorphic threats. Rule-based firewalls require manual updates and fail against novel attack vectors. The shared responsibility model of cloud security creates ambiguity about which party is accountable for specific protections (sharma2024cloud?). These limitations collectively motivate AI-driven, adaptive defense architectures.

III. AI-BASED THREAT DETECTION AND CLASSIFICATION

A. Overview

The application of machine learning and deep learning to threat detection represents the most mature and extensively studied area in AI-driven cloud security. Surveyed works demonstrate a clear evolution from static, rule-based approaches toward adaptive, data-driven models capable of handling high-dimensional, high-velocity telemetry data.

B. Deep Learning for DDoS Defense

Afraji et al. (afraji2025ddos?) provide a comprehensive review of deep learning (DL) models applied to DDoS detection in cloud and IoT environments. The authors categorize DDoS attacks into volumetric (flooding bandwidth), protocol-based (exploiting stack vulnerabilities), and application-layer (mimicking legitimate traffic) variants. They evaluate Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks, noting that deep learning significantly outperforms traditional algorithms in detecting high-volume attack patterns. Critically, the paper identifies a persistent limitation: detection performance is heavily constrained by the quality and diversity of available datasets. The authors advocate for greater use of Explainable AI (XAI) to improve trust in automated detection decisions, particularly in production environments.

C. Hybrid GNN-Transformer Architecture for Cloud Threat Detection

Nutalapati et al. (nutalapati2025cloud?) propose an AI-driven threat detection framework for cloud environments that combines Graph Neural Networks (GNNs) with transformer-based attention mechanisms. GNNs model the relational dependencies between cloud entities — virtual machines, containers, microservices — enabling detection of lateral movement and privilege escalation patterns. Transformer attention heads capture temporal and contextual correlations across event sequences, enabling zero-day threat identification through behavioral deviation rather than pattern matching.

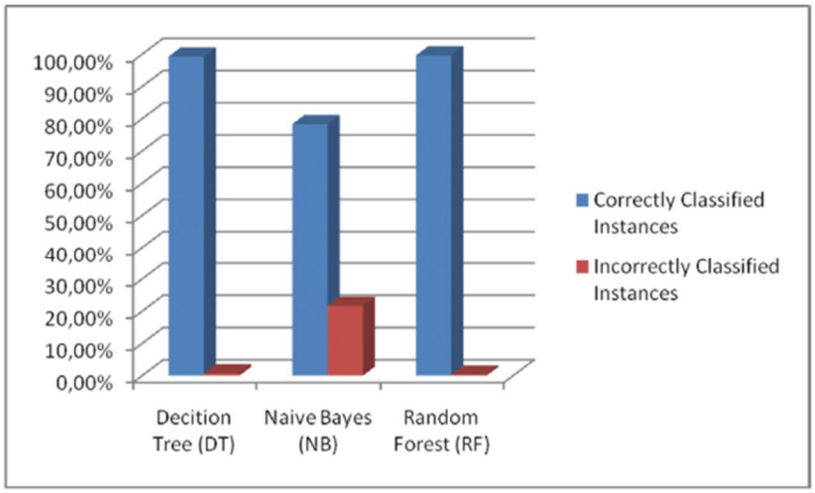
The system additionally incorporates self-supervised contrastive learning to detect never-before-seen attack variants using only limited labeled data, and a reinforcement learning (RL)-based autonomous mitigation engine. Evaluated on the UNSW-NB15 dataset, the framework achieves 97.8% accuracy, 97.2% precision, 97.5% recall, and 97.3% F1-score, representing state-of-the-art performance.

D. Federated Deep Learning for Ransomware and Malware Detection

Prajwalasimha et al. (prajwala2025ransomware?) address the detection of ransomware and malware in distributed edge-cloud environments, where centralized training models are impractical due to privacy constraints, latency, and scalability. Their framework leverages Federated Deep Learning (FDL), allowing edge devices to train local models without exposing raw data to central servers — a critical advantage for GDPR compliance. The architecture integrates GNNs to model relationships between system entities, transformer encoders to capture temporal evolution of those relationships, and a multi-layer threat intelligence fusion engine that combines system call graphs, encrypted network traffic analysis, and process lineage information. Evaluated on hybrid real-world datasets, the system achieves 98.7% detection accuracy, reduces detection latency by 46%, and decreases false alarm rates by 62% compared to state-of-the-art baselines.

E. Classifier Performance: Random Forest vs. Alternatives

A recurring theme across surveyed works is the choice of classification algorithm. Figure 1 illustrates the comparative classification accuracy of three widely used algorithms — Decision Tree (DT), Naive Bayes (NB), and Random Forest (RF) — on intrusion detection datasets. Both DT and RF achieve near-perfect classification rates, while NB exhibits significantly higher misclassification, underscoring why ensemble methods such as Random Forest are preferred for high-stakes, low-tolerance security applications (swetha2025ai?; mishra2025ai?). Random Forest's robustness stems from its ensemble of independent decision trees, which collectively reduce variance and overfitting — critical properties when classifying noisy, high-dimensional network telemetry.



Classification accuracy comparison of Decision Tree (DT), Naive Bayes (NB), and Random Forest (RF) on intrusion detection data. RF achieves near-perfect accuracy with minimal misclassification, validating its selection as the primary classifier in the proposed system.

F. AI-Based Threat Detection: A General Framework

Mishra et al. (mishra2025ai?) present a general AI-based cybersecurity framework emphasizing deep learning and reinforcement learning for threat detection and response. Their model is designed around four key capabilities: anomaly detection across large datasets, attack pattern recognition, automated response through policy enforcement, and behavioral analytics for insider threat identification. The authors highlight the inadequacy of signature-based systems against polymorphic and zero-day threats, and demonstrate that integrating RL into security decision pipelines enables adaptive, context-aware responses. The framework explicitly addresses ethical standards including transparency and bias mitigation in AI-driven security decisions.

G. *Comprehensive AI Review for Cybersecurity*

Swetha et al. (swetha2025ai?) provide a broad survey of AI techniques applied to cybersecurity, covering ML and DL methods across intrusion detection, malware analysis, phishing detection, and user authentication. They introduce the Q-PERAL methodology (Quantum-enhanced Predictive, Explainable, Resilient, Adaptive Learning) as a conceptual framework for next-generation cybersecurity AI. Notably, the survey distinguishes between offensive AI (automated hacking frameworks such as DeepLocker and AutoSploit, adversarial machine learning) and defensive AI (anomaly detection, adversarial training, real-time threat analysis), framing the security problem as an adversarial arms race that demands continuous model evolution.

IV. EXPLAINABLE AI (XAI) FOR CYBERSECURITY

A. *The Black Box Problem in Cybersecurity AI*

A fundamental challenge in deploying AI for cybersecurity is the opacity of high-performing models. Deep neural networks, ensemble methods, and other complex architectures achieve strong detection accuracy but function as “black boxes” — their internal decision processes are not human-interpretable. In security contexts, this opacity is particularly problematic: security analysts must understand *why* a system flagged an activity as malicious before taking consequential actions such as blocking traffic or isolating infrastructure (sharma2025xai?; afraji2025ddos?).

B. *Taxonomy of XAI Techniques*

Sharma et al. (sharma2025xai?) provide a comprehensive survey of XAI techniques applied to cybersecurity, presenting a structured taxonomy across multiple dimensions. The authors distinguish between:

Model-agnostic methods: Techniques applicable to any underlying model, including LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations). These generate post-hoc explanations by approximating model behavior locally or by computing feature contribution scores across predictions.

Model-specific methods: Intrinsic interpretability tools tied to particular architectures, such as attention visualization for transformer models or decision path extraction for tree-based ensembles.

Ante-hoc vs. post-hoc: Ante-hoc methods build interpretability directly into model design (e.g., decision trees, linear models), while post-hoc methods explain the outputs of pre-trained black-box models.

The survey applies this taxonomy across three major threat categories — malware detection, phishing identification, and network intrusion detection — demonstrating how XAI can surface actionable insights for each. The paper also reviews evaluation criteria for XAI effectiveness, including fidelity, stability, comprehensibility, and completeness.

C. *XAI for DDoS Defense*

Afraji et al. (afraji2025ddos?) separately emphasize XAI as a necessary complement to deep learning-based DDoS defense. The authors argue that without explainability, security operators cannot verify whether detection models are learning genuine attack signatures or spurious correlations. They call for integrated XAI pipelines that can provide real-time explanations alongside detection outputs, enabling faster and more confident analyst decision-making.

D. *Implications for Autonomous Systems*

Both papers converge on a critical insight for autonomous defense systems: explainability is not merely a transparency feature but a prerequisite for operational trust and legal compliance. Autonomous systems that act on model outputs without interpretable justification may violate audit requirements and undermine incident response workflows. Future systems must couple high-accuracy detection with human-readable explanations of each decision.

V. AUTONOMOUS THREAT RESPONSE AND MITIGATION

A. *From Detection to Autonomous Action*

Detecting a threat is necessary but insufficient — effective security requires timely, appropriate response. Autonomous response systems close the loop between detection and mitigation, executing protective actions without waiting for human intervention. The surveyed literature explores several architectural approaches to autonomous response in cloud environments.

B. Reinforcement Learning for Adaptive Response

Nutalapati et al. (nutalapati2025cloud?) integrate a reinforcement learning-based autonomous mitigation engine into their cloud security framework. The RL agent learns to balance competing objectives: containing detected threats while maintaining service availability. By calculating trade-offs through reward modeling, the agent selects from a repertoire of response actions — IP blocking, container isolation, traffic throttling, firewall rule modification — and adapts its policy based on the outcome of previous interventions.

C. Multi-Cloud AI Security Management

Srimathi et al. (srimathi2023multicloud?) address the particular complexity of hybrid multi-cloud environments, where security policies, data residency requirements, and vendor-specific controls must be coordinated across multiple providers. Their framework proposes AI-driven dynamic threat assessment to prioritize vulnerabilities, reducing the cognitive load on security teams and enabling faster triage. A mathematical model for optimal resource allocation is presented, ensuring that security compute resources are efficiently distributed across cloud workloads. The paper highlights the role of predictive AI in anticipating attack vectors before they are exploited, enabling proactive rather than reactive security postures.

D. Cloud Security Optimization

Sharma and Singla (sharma2024cloud?) propose a combined security model integrating Identity and Access Management (IAM), data encryption, and AI/ML-enhanced network security for organizational cloud deployments. Their empirical evaluation demonstrates an Incident Mitigation Rate of 92%, Response Time Reduction of 85%, Resource Utilization Efficiency of 78%, and User Satisfaction of 88% — all exceeding benchmark models. These metrics validate that integrated, AI-assisted security frameworks substantially outperform siloed, traditional approaches.

E. Cybersecurity in Virtualized Environments

Krishnappa et al. (krishnappa2024cloud?) examine the specific security challenges of virtualized cloud environments, including hypervisor vulnerabilities, VM escape attacks, and container isolation failures. Their proposed layered security framework combines encryption, IAM, intrusion detection, and AI-driven threat detection with quantum-safe cryptography for forward-looking resilience. Case studies and benchmark analyses validate the framework's effectiveness in improving data confidentiality, integrity, and availability across shared cloud infrastructure.

F. General AI Framework for Threat Mitigation

Mishra et al. (mishra2025ai?) describe behavioral analytics as a key mechanism for insider threat detection within autonomous response pipelines. Their system monitors user activity at granular levels, flags anomalies in real time, and applies automated isolation or access revocation when suspicious behavior is detected. This approach demonstrates that autonomous response need not be limited to network-level interventions — it can extend to identity and access management layers, enabling comprehensive defense-in-depth.

VI. REAL-TIME CYBER THREAT ATTRIBUTION

A. The Attribution Problem

Attributing cyberattacks to their true sources is a prerequisite for effective deterrence, legal action, and strategic response. However, attribution is complicated by attacker use of proxies, VPNs, botnet infrastructure, spoofed IP addresses, and malware signatures that deliberately mimic other actors (mangaiyarkarasi2025attribution?). Traditional attribution relies on static forensics that are slow, labor-intensive, and easily circumvented.

B. Dynamic Multi-Mode Source Attribution (DMSA)

Mangaiyarkarasi et al. (mangaiyarkarasi2025attribution?) propose the Dynamic Multi-Mode Source Attribution (DMSA) architecture, a meta-learning-based AI system designed for adaptive, explainable, real-time cyber attribution. DMSA deploys multiple self-learning AI agents that compete and collaborate to improve attribution accuracy and resilience. Key innovations include:

1) Multi-modal telemetry fusion: DMSA aggregates evidence from endpoints, networks, cloud services, and dark web sources, constructing a unified evidence graph that integrates signals across heterogeneous data streams.

2) Dynamic confidence scoring: Rather than producing binary attributions, DMSA outputs probabilistic confidence scores that reflect uncertainty and enable analysts to calibrate their response accordingly.
3) Adversarial resilience: The system retests attribution decisions when new or conflicting evidence arrives, maintaining robustness against camouflaged attacks and deliberate misdirection by sophisticated actors.

Experimental results demonstrate that DMSA substantially reduces attribution latency and improves accuracy compared to traditional rule-based and static signature methods, representing a meaningful advancement in scalable AI-based cyber attribution.

C. Relevance to Autonomous Defense

Real-time attribution enhances autonomous response by providing contextual intelligence about attack origin, sophistication, and likely intent. An autonomous defense system equipped with attribution capabilities can tailor response strategies — for example, applying stricter isolation for nation-state-level threats versus opportunistic automated attacks — improving both the precision and proportionality of automated interventions.

VII. AI-ENHANCED CYBERSECURITY EDUCATION

A. The Workforce Gap

Effective deployment of AI-driven security systems requires a skilled human workforce capable of configuring, monitoring, and improving autonomous defenses. However, the global cybersecurity talent shortage — driven in part by static educational curricula that lag behind evolving threats — represents a critical vulnerability in the overall security ecosystem (**tolah2026teachsecure?**). Nearly three-quarters of cybersecurity professionals report gaps between their academic training and the demands of real-world roles.

B. TeachSecure-CTI: Adaptive Curriculum Generation

Tolah (tolah2026teachsecure?) introduces TeachSecure-CTI, a framework for adaptive cybersecurity curriculum generation that integrates real-time Cyber Threat Intelligence (CTI) with AI-driven personalization. The system architecture comprises three main components:

1) CTI ingestion and clustering: Real-time threat intelligence feeds are ingested, clustered by attack category, and mapped to MITRE ATT&CK tactic taxonomy to ensure curriculum content reflects current adversarial behaviors.
2) NLP-based semantic concept extraction: Natural language processing pipelines extract key concepts from CTI reports and learning materials, enabling semantic alignment between threat content and instructional modules.
3) Reinforcement learning for content sequencing: An RL agent adaptively sequences learning content based on individual learner performance profiles and the evolving threat landscape, optimizing for competency development velocity.

A 12-week study across 150 students at three institutions demonstrated that TeachSecure-CTI improves learning gains by 34% over static curricula, achieves 84.8% personalization accuracy, and delivers 85.9% recognition accuracy for MITRE ATT&CK tactics, with a 31% faster competency development rate.

C. Implications

The AI techniques underlying TeachSecure-CTI — RL-based personalization, CTI integration, NLP-driven content alignment — directly mirror those used in autonomous defense systems. The paper illustrates how the same AI infrastructure that powers threat detection can be repurposed to accelerate human expertise development, creating a positive feedback loop between autonomous systems and the human operators who oversee them.

VIII. COMPARATIVE ANALYSIS

Table [tab:comparison] summarizes the key characteristics, techniques, and contributions of all surveyed papers across the five thematic areas.

Paper	Theme	Key Techniques	Dataset / Evaluation	Key Result
Nutalapati et al. (nutalapati2025cloud?)	Detection & Response	GNN, Transformer, RL, Self-supervised learning	UNSW-NB15	97.8% accuracy
Prajwalasimha et al.	Detection	FDL, GNN,	Hybrid real-world	98.7% detection;

Paper	Theme	Key Techniques	Dataset / Evaluation	Key Result
(prajwala2025ransomware?)		Transformer, Threat intelligence fusion		46% latency reduction
Afraji et al. (afraji2025ddos?)	Detection, XAI	CNN, RNN, LSTM, XAI	Survey / review	XAI necessity highlighted
Mishra et al. (mishra2025ai?)	Detection & Response	Deep learning, RL, Behavioral analytics	Conceptual framework	Scalable, adaptive defense
Swetha et al. (swetha2025ai?)	Detection (survey)	ML, DL, Q-PERAL methodology	Survey / review	Offensive vs. defensive AI taxonomy
Sharma et al. (sharma2025xai?)	XAI	LIME, SHAP, Attention, Model-agnostic	Survey / review	XAI taxonomy for cybersecurity
Srimathi et al. (srimathi2023multicloud?)	Response (multi-cloud)	AI threat assessment, Resource optimization	Conceptual + real-world	Proactive multi-cloud security
Sharma & Singla (sharma2024cloud?)	Response	IAM, Encryption, AI/ML integration	Organizational deployment	92% mitigation rate
Krishnappa et al. (krishnappa2024cloud?)	Response	Layered defense, AI detection, Quantum crypto	Case studies	Improved CIA triad metrics
Akbari et al. (akbari2024cloud?)	Cloud security (survey)	Multi-layered security, AI, MFA	Survey / review	Cloud security taxonomy
Mangaiyarkarasi et al. (mangaiyarkarasi2025attribution?)	Attribution	Meta-learning, Multi-modal fusion, RL agents	Experimental	Reduced attribution latency
Tolah (tolah2026teachsecure?)	Education	RL, NLP, CTI integration, MITRE ATT&CK	150 students, 3 institutions	34% learning gain improvement

A. Key Observations

1) Algorithm convergence: GNNs, transformer attention mechanisms, and reinforcement learning appear across multiple independent works (nutalapati2025cloud?; prajwala2025ransomware?; mishra2025ai?), suggesting these are emerging as foundational primitives for cloud security AI.

2) Dataset dependency: Most empirical papers rely on benchmark datasets (UNSW-NB15, NSL-KDD, CICIDS2017). While these enable reproducibility, they may not fully capture the distribution of attacks in modern cloud workloads, limiting generalization claims (afraji2025ddos?; swetha2025ai?).

3) XAI as a cross-cutting concern: Explainability is identified as a requirement across detection, response, and attribution tasks, yet few systems provide native XAI integration alongside their primary detection or response capabilities (sharma2025xai?).

4) Privacy-preserving learning: Federated learning emerges as a key enabler for privacy-compliant, distributed cloud security (prajwala2025ransomware?), particularly relevant under GDPR and similar regulatory frameworks.

5) Closed-loop architectures: The most effective systems surveyed implement feedback mechanisms that use response outcomes to improve future detection and decision-making (nutalapati2025cloud?; mishra2025ai?; srimathi2023multicloud?).

IX. OPEN CHALLENGES AND FUTURE DIRECTIONS

A. Dataset Quality and Diversity

A recurring limitation across surveyed works is dependence on benchmark datasets that do not reflect the heterogeneity and dynamism of real-world cloud workloads (afraji2025ddos?; swetha2025ai?). Future work should prioritize the creation of large-scale, continuously updated cloud-native datasets that capture multi-tenant traffic, containerized workloads, and contemporary attack campaigns.

B. Adversarial Robustness

AI-driven security systems are themselves vulnerable to adversarial manipulation. Attackers can craft inputs specifically designed to evade ML-based detection (swetha2025ai?; sharma2025xai?). Building models that are robust to adversarial perturbation — particularly in real-time, high-stakes environments — remains an open research challenge.

C. Scalability and Latency

Cloud environments operate at massive scale with stringent latency requirements. Many proposed architectures demonstrate strong performance on benchmark datasets but have not been validated at production scale (nutalapati2025cloud?; prajwala2025ransomware?). Efficient model inference, edge deployment, and federated training are key areas for continued investigation.

D. Interoperability and Standards

Multi-cloud and hybrid cloud deployments involve heterogeneous security tools, APIs, and data formats from multiple vendors (srimathi2023multicloud?; akbari2024cloud?). Autonomous defense systems must operate effectively across these boundaries, motivating standardized threat intelligence sharing protocols and vendor-agnostic security APIs.

E. Ethical and Legal Considerations

Autonomous security systems that take consequential actions — blocking traffic, isolating services, modifying access policies — must operate within ethical and legal frameworks (mishra2025ai?; sharma2025xai?). Issues of algorithmic bias, accountability for false positives, and compliance with data protection regulations (GDPR, CCPA) require careful consideration in system design.

F. Human-AI Collaboration

While autonomous systems reduce the burden on human operators, they do not eliminate the need for human judgment in complex or novel situations. Designing effective human-AI interfaces that convey system reasoning, support override mechanisms, and maintain operator situational awareness is an underexplored but critical challenge (sharma2025xai?; tolah2026teachsecure?).

X. CONCLUSION

This survey has systematically reviewed thirteen recent papers to characterize the state of the art in AI-driven cybersecurity for cloud environments. We have identified five core research themes — threat detection, explainable AI, autonomous response, real-time attribution, and cybersecurity education — and synthesized the key techniques, findings, and limitations across each.

The literature converges on several important conclusions. First, deep learning architectures combining GNNs, transformer attention, and reinforcement learning represent the current frontier of autonomous cloud security, achieving detection accuracies exceeding 97% on benchmark datasets. Second, explainability is a non-negotiable requirement for real-world deployment of AI security systems, demanding integration of XAI techniques alongside primary detection capabilities. Third, federated learning provides a privacy-preserving path to distributed model training across cloud tenants and edge devices. Fourth, closed-loop systems that couple detection with autonomous response and feedback-driven learning substantially outperform static, reactive approaches.

Significant challenges remain, including dataset limitations, adversarial robustness, production-scale validation, and the legal and ethical governance of autonomous security actions. Addressing these challenges will require interdisciplinary collaboration across machine learning, systems engineering, security policy, and human-computer interaction.

The surveyed works collectively demonstrate that AI-driven, autonomous approaches to cloud cyber defense are not only technically feasible but increasingly necessary given the scale and sophistication of modern threats. The convergence of ensemble learning, deep neural architectures, reinforcement learning, and explainability techniques points toward a future in which cloud environments are protected by intelligent, self-adapting systems capable of detecting, attributing, and mitigating threats at machine speed — with minimal human intervention and maximum transparency.

REFERENCES

[1] Tolah, “TeachSecure-CTI: Adaptive Cybersecurity Curriculum Generation Using Threat Dynamics and AI,” Computers, Materials & Continua, vol. 87, no. 1, pp. 71–[end], 2026. DOI: 10.32604/cmc.2025.074997.

[2] Sharma, S. Rani, and M. Shabaz, “A comprehensive review of explainable AI in cybersecurity: Decoding the black box,” ICT Express, vol. 11, pp. 1200–1219, 2025. DOI: 10.1016/j.icte.2025.10.004.

[3] D. M. A. A. Afraji, J. Lloret, and L. Peñalver, "Deep learning-driven defense strategies for mitigating DDoS attacks in cloud computing environments," Cyber Security and Applications, vol. 3, p. 100085, 2025. DOI: 10.1016/j.csa.2025.100085.

[4] T. Swetha, U. Kumaran, V. P. Meena, and I. A. Hameed, "Leveraging AI for enhanced cybersecurity: a comprehensive review," Discover Applied Sciences, vol. 7, p. 584, 2025. DOI: 10.1007/s42452-025-06773-0.

[5] R. Sharma and A. Singla, "Optimizing Cloud Security: A Study of Effective Cybersecurity Measures for Organizations," in Proc. International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS-2024), IEEE, 2024. DOI: 10.1109/ICICNIS.2024.10823237.

[6] V. Mangaiyarkarasi et al., "Real-Time Source Attribution of Cyberattacks via AI-Driven Threat Intelligence," in Proc. 2025 International Conference on Innovations and Emerging Technologies In AI & Communication Systems (IETACS), IEEE, 2025. DOI: 10.1109/IETACS.2025.10786533.

[7] M. Mishra, R. R. Pradhan, K. Agrawalla, and R. Bokka, "AI for Cybersecurity Threat Detection: A Machine Enabled Computing Perspective," in Proc. 2025 International Conference on Smart, Secure, Intelligent Advances Computing (SISSAC), IEEE, 2025. DOI: 10.1109/SISSAC.2025.11158640.

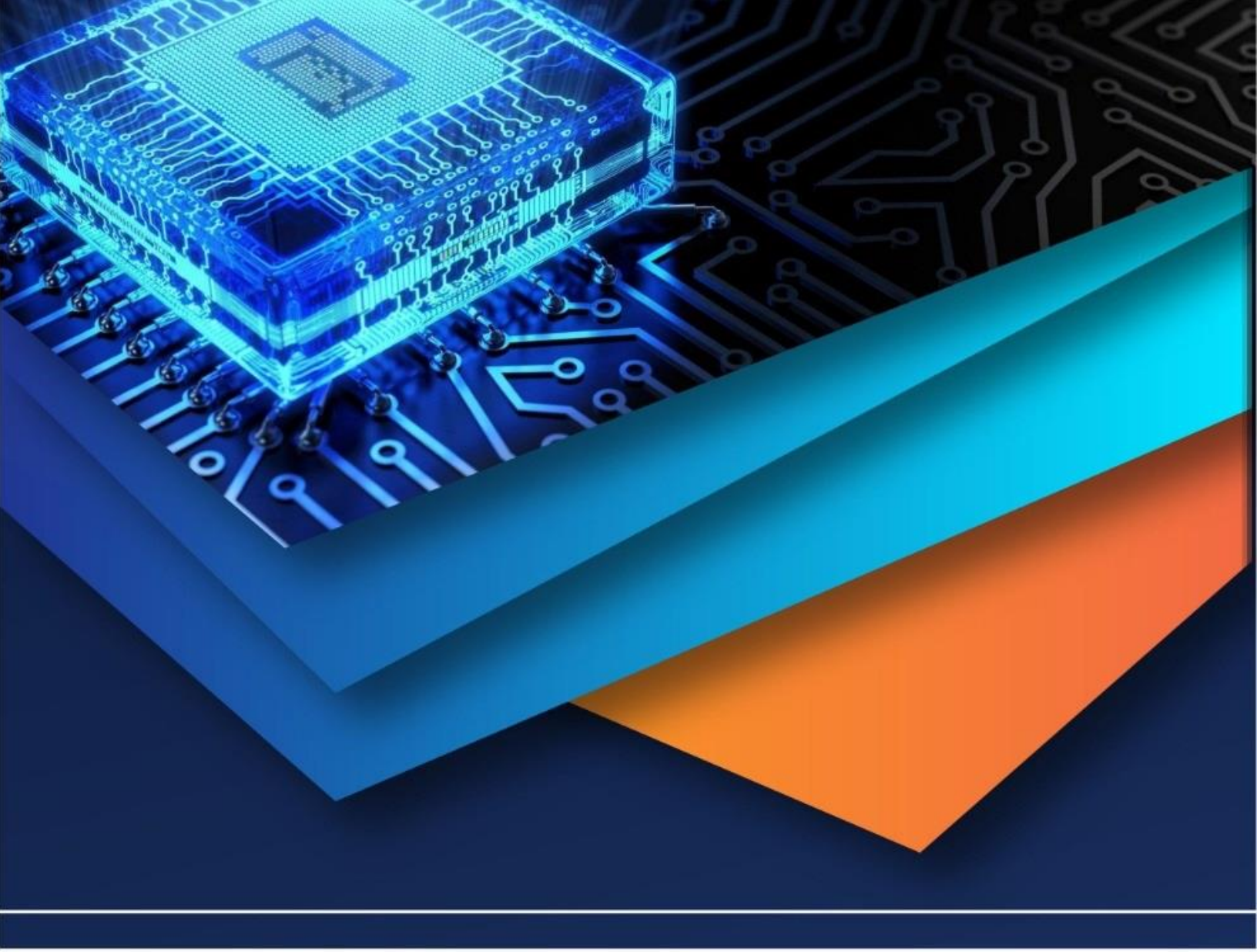
[8] P. S. N. Prajwalasimha et al., "AI-Powered Predictive Ransomware and Malware Detection for Edge and Cloud Systems," in Proc. 2025 International Conference on NexGen Networks and Cybernetics (IC2NC), IEEE, 2025. DOI: 10.1109/IC2NC.2025.10367310.

[9] M. S. Krishnappa et al., "Cybersecurity in the Cloud Era: Protecting Virtualized Environments Against Evolving Threats," in Proc. 2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA), IEEE, 2024. DOI: 10.1109/ICICyTA.2024.10190113.

[10] P. Nutalapati, A. K. Elengovan, N. Prakash, I. Sahoo, and V. Nutalapati, "AI-Driven Threat Detection in Cloud Environment: Towards Autonomous Security Systems," in Proc. 10th International Conference on Communication and Electronics Systems (ICCES-2025), IEEE, 2025. DOI: 10.1109/ICCES.2025.10378836.

[11] M. Akbari, R. Bruschi, and A. Carrega, "Navigating the Cybersecurity Landscape in Cloud Computing: Challenges, Strategies, and Future Directions," University of Genoa, 2024.

[12] J. Srimathi et al., "AI-Enhanced Multi-Cloud Security Management: Ensuring Robust Cybersecurity in Hybrid Cloud Environments," in Proc. 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), IEEE, 2023. DOI: 10.1109/ICSES.2023.10055640.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY