ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may1257

Explainable AI for Memory Artifact Triage in Serverless Cloud Forensics

Onyagu Chika Lilian[1]; Ifeanyichukwu Oluchukwu Aniakor[2]; Obidike Chukwuemeka Augustine[3]; Adigwe Amaka Raechel[4]; Edeh Hyacinth[5]

[1]Department of Cybersecurity and Data Science
[2]Department of Computer Science, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka Nigeria.
[3]Department of Computer Science, Delta State University, Abraka, Delta State, Nigeria

Publication Date: 2026/06/03

Abstract: Serverless cloud computing is becoming increasingly popular due to its scalability, flexibility, and ease of infrastructure administration. Platforms such as AWS Lambda and Knative support event-driven execution of applications without direct server administration. However, the transitory and distributed nature of serverless settings poses substantial hurdles to digital forensics, notably in the acquisition, retention, and analysis of volatile memory artifacts. Traditional forensic techniques are often ineffective because critical evidence may disappear rapidly, while the large volume of generated artifacts complicates incident investigation and response. Furthermore, many AI-based forensic systems use black-box models, which reduces transparency and trust in forensic decision-making. This study presents an Explainable Artificial Intelligence (XAI)-based system for memory artifact triage in serverless cloud computing settings. The framework integrates Graph Neural Networks (GNNs) for analyzing relationships among forensic artifacts, explainability techniques such as SHAP, LIME, and GNNExplainer for interpretable decision-making, and Large Language Models (LLMs) for generating human-readable forensic explanations. Memory artifacts including processes, API calls, execution traces, and network interactions, are represented as graph structures to support anomaly detection, artifact classification, and suspicious behavior identification. The explainability layer reveals the reasoning behind forensic choices, enhancing accountability and forensic validation. The suggested framework improves forensic readiness, evidence prioritization, transparency, and incident response efficiency in cloud-native systems, while adhering to standards like ISO/IEC 27037. The report also highlighted issues with scalability, privacy, and the ethical implications of auditable AI in digital investigations.

Keywords: *Serverless Forensics; Explainable AI; Memory Triage; Graph Neural Networks; Cloud Security.*

How to Cite: Onyagu Chika Lilian; Ifeanyichukwu Oluchukwu Aniakor; Obidike Chukwuemeka Augustine; Adigwe Amaka Raechel; Edeh Hyacinth (2026) Explainable AI for Memory Artifact Triage in Serverless Cloud Forensics.
International Journal of Innovative Science and Research Technology, 11(5), 2980-2984.
https://doi.org/10.38124/ijisrt/26may1257

I. INTRODUCTION

Cloud computing is rapidly advancing as a foundational technology, offering robust prospects for enterprises in the foreseeable future [1]. However, serverless computing encounters performance challenges stemming from infrastructure variability and the scheduling policies enforced by cloud service providers [2]. Based on deployment strategies, cloud computing is categorized into public, private, hybrid, and community models. Service delivery is further segmented into Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [1]. A defining feature of modern cloud environments is serverless computing, which enables developers to build and deploy applications without managing the underlying infrastructure. In these systems, cloud providers dynamically allocate resources and execute tasks on demand, with Google Cloud Functions and AWS Lambda serving as prominent examples.

While serverless architecture delivers cost efficiency, scalability, and flexibility, it introduces formidable obstacles for digital forensics. The sheer volume of data and the ephemeral nature of resources are the primary hurdles. Serverless forensics entails the identification, collection, preservation, analysis, and presentation of digital evidence from these environments. Traditional forensic methodologies frequently falter here because resources are short-lived, highly distributed, and exceptionally dynamic. Consequently, forensic investigators struggle to secure reliable evidence while maintaining data integrity and regulatory compliance as organizations increasingly adopt serverless solutions. Fig. 1 illustrates the life cycle of AWS Lambda, highlighting the transient nature of function execution.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may1257

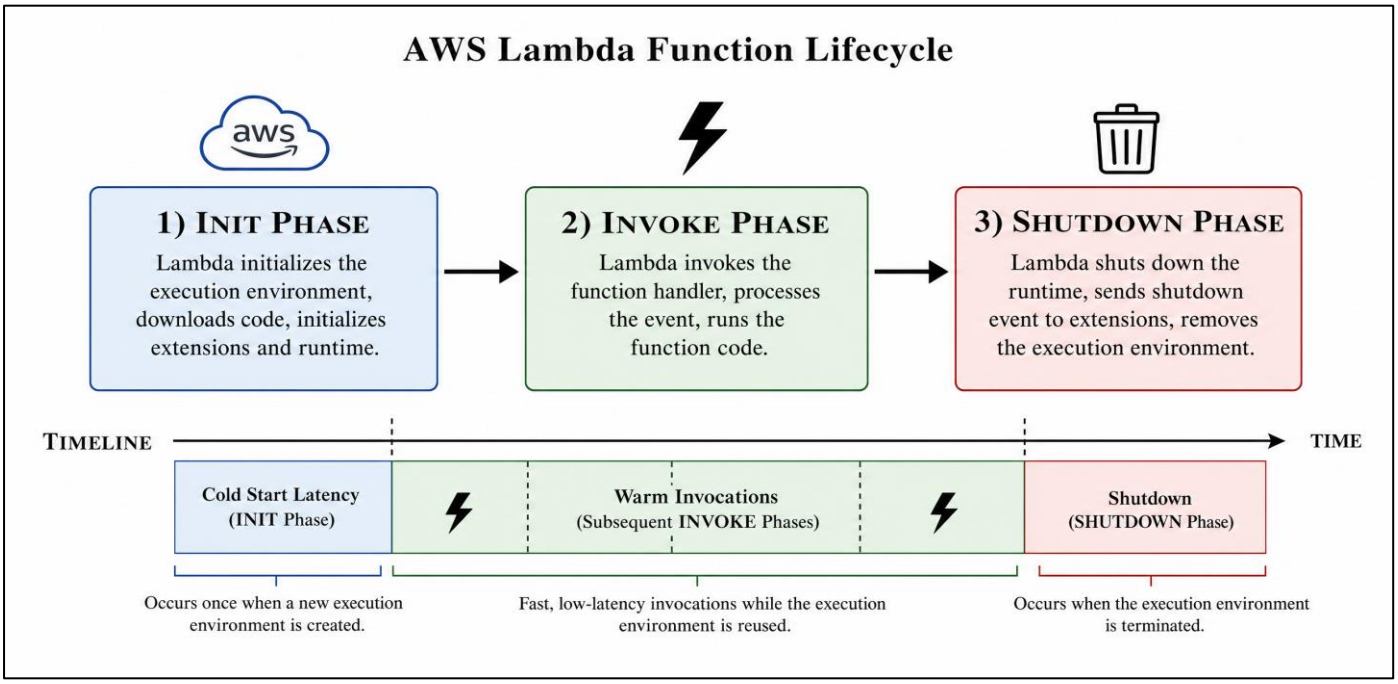

Fig 1 AWS Lambda Function Lifecycle

There is a critical demand for Explainable Artificial Intelligence (XAI) within cloud security to supply not only robust detection mechanisms but also transparent, interpretable rationales for automated decisions [3]. XAI encompasses techniques designed to render the operations of artificial intelligence systems comprehensible to human analysts. The objective is to foster transparency, trust, accountability, and interpretability in machine learning models. As AI becomes deeply embedded in cloud-native and serverless ecosystems, a significant gap emerges between the necessity for XAI and the inherent traits of serverless computing. The lack of transparency, ephemerality, scalability constraints, fragmented logging, and heightened security threats impede the effective deployment of responsible AI in these settings.

This study aims to develop an XAI-driven framework tailored for memory artifact triage in serverless cloud environments. The goal is to refine digital forensic investigations, enhance decision-making transparency, and mitigate the challenges posed by ephemeral infrastructures. Specifically, the research investigates the difficulties of acquiring volatile memory artifacts, categorizes evidence generated during serverless execution, and constructs an intelligent triage mechanism to prioritize relevant data. By embedding explainable AI, investigators can decipher the rationale behind artifact classification. The framework's efficacy is evaluated using metrics such as accuracy, precision, recall, F1-score, and response time, contrasting its performance with traditional black-box AI models regarding transparency and forensic readiness.

This research uniquely intersects XAI with memory artifact triage in serverless contexts—a domain scarcely explored in prior digital forensic literature. Unlike conventional systems reliant on manual analysis or opaque algorithms, the proposed interpretable AI-driven method offers accessible explanations for forensic determinations. Furthermore, it introduces an intelligent prioritization system that streamlines investigation complexity and accelerates response times in highly dynamic cloud environments.

II. RELATED LITERATURE

The acquisition and analysis of volatile memory to detect cyber threats remain a dynamic focus within cybersecurity research. As computer systems increasingly underpin modern society, malicious actors continually devise innovative methods to breach them [4]. In digital forensics, law enforcement and institutional investigators deploy various commercial and open-source tools to scrutinize digital evidence. These tools are indispensable for addressing contemporary security challenges across computer, network, mobile, cloud, and IoT forensics [5]. Fileless malware, which leverages legitimate programs to infect systems without leaving traces on the hard drive, is proliferating and frequently evades standard antivirus defenses [4].

Memory forensics has evolved from examining volatile memory on isolated machines to conducting sophisticated investigations within virtualized, cloud, and serverless environments. Initially, the focus was on extracting Random Access Memory (RAM) from physical hardware to recover transient data—such as running processes, encryption keys, network connections, and malware artifacts—that do not persist on disk. While early investigations leaned heavily on disk forensics, it became evident that crucial evidence often resides exclusively in memory during system operation. Specialized frameworks like Volatility and Rekall have significantly improved the detection of advanced malware and rootkits. With the advent of cloud computing, memory forensics expanded to encompass distributed infrastructures, introducing challenges like multi-tenancy and restricted hardware access. The rise of containerization and serverless

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may1257

computing has further revolutionized the field, creating ephemeral systems where memory artifacts vanish almost instantly. Consequently, modern memory forensics increasingly relies on automation, machine learning, and XAI to enhance artifact triage, threat detection, and transparency.

Malware—software designed to infiltrate systems covertly and exploit resources—remains a pervasive global threat [6]. It continually evolves, employing sophisticated techniques to bypass traditional detection and exploit vulnerabilities [7]. This proliferation necessitates innovative, adaptive mitigation strategies, as signature-based approaches often fail against zero-day and polymorphic attacks [8]. Malware analysis is generally bifurcated into static and dynamic methods. Static analysis inspects binaries without execution through fingerprinting and disassembling, though it struggles with obfuscated code. Conversely, dynamic analysis observes the malware's behavior during execution in a controlled environment [6].

XAI has gained prominence in malware analysis because many contemporary deep learning models function as black boxes. Techniques like SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) help analysts understand AI decisions by highlighting the features that drive predictions. SHAP utilizes cooperative game theory to assign relevance values to specific features, offering both local and global interpretability. LIME explains individual predictions by approximating complex model behaviors with simpler, interpretable models. In practice, these tools reveal whether API calls, network activity, or memory access patterns influenced the classification of a process as malicious or benign. Integrating SHAP and LIME enhances transparency, allowing analysts to validate AI judgments, identify false positives, and foster evidence-based cybersecurity responses.

Serverless computing fundamentally alters cloud architecture, enabling developers to focus solely on code deployment without managing infrastructure [9]. Cloud providers automatically handle resource provisioning, load balancing, and execution. Applications are typically deployed as discrete functions triggered by events such as HTTP requests or database updates, adhering to the Function-as-a-Service (FaaS) model where billing is based strictly on execution time. This paradigm boosts scalability and cost efficiency while reducing operational complexity. AWS Lambda automatically scales applications by executing multiple function instances concurrently. It integrates seamlessly with services like Amazon S3 and DynamoDB, making it ideal for real-time processing and microservices. However, the short-lived, stateless nature of Lambda functions exacerbates the ephemerality challenges in serverless forensics.

Knative is an open-source, Kubernetes-based framework that facilitates serverless workloads across hybrid and multi-cloud environments, offering organizations greater control over their infrastructure. It comprises Knative Serving for workload deployment and Knative Eventing for event-driven communication. A key feature is scale-to-zero, allowing idle programs to shut down and restart upon receiving new requests. While AWS Lambda provides a fully managed, proprietary experience, Knative offers portability and vendor independence. Despite their distinct advantages, both platforms' highly distributed and ephemeral execution environments present significant hurdles for monitoring, security, memory artifact acquisition, and digital forensics.

III. METHODOLOGY

The study synergizes Graph Neural Networks (GNNs) and Large Language Models (LLMs) to forge an intelligent, explainable framework for memory artifact triage in serverless cloud environments. This framework is engineered to parse complex relationships among memory artifacts via graph-based learning, while simultaneously generating comprehensible forensic explanations using natural language processing. Fig. 2 illustrates the overall pipeline of the proposed framework.

- *Data Acquisition*

The initial phase involves harvesting volatile memory artifacts from serverless platforms like AWS Lambda and Knative. Given the ephemeral nature of these functions, data collection must be executed in near real-time. Acquired artifacts encompass running processes, API calls, system calls, network connections, file access events, memory allocation patterns, container metadata, authentication tokens, function invocation logs, synthetic serverless dumps, and AWS traces. This evidence is temporarily secured in centralized forensic storage for subsequent analysis.

- *Graph Neural Network (GNN) Analysis Layer*

The constructed artifact graph is ingested by a Graph Neural Network model designed to detect anomalous patterns and prioritize forensic artifacts based on their relational context. Nodes represent individual artifacts while edges encode the relationships between them, enabling the model to identify suspicious interaction patterns that would be invisible in flat, non-relational data structures.

- *Explainability Layer*

To bolster transparency, explainability mechanisms such as SHAP, LIME, and GNNExplainer are woven into the GNN analysis workflow, illuminating the features driving the model's classifications. This layer is critical for ensuring that forensic investigators can audit and validate the AI's reasoning, which is a prerequisite for legal admissibility.

- *Large Language Model (LLM) Explanation Layer*

The outputs from the GNN and the explainability module are routed to a Large Language Model (LLM), which translates the technical findings into natural language interpretations suitable for human analysts. This layer bridges the gap between complex AI outputs and actionable forensic intelligence.

- *Decision and Triage Layer*

In the final stage, forensic artifacts are prioritized based on their relevance and associated threat levels. The system categorizes artifacts into tiers: Critical, High Risk, Moderate

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may1257

Risk, and Benign. This stratification allows investigators to concentrate on the most pertinent evidence immediately, thereby reducing investigation time and complexity.

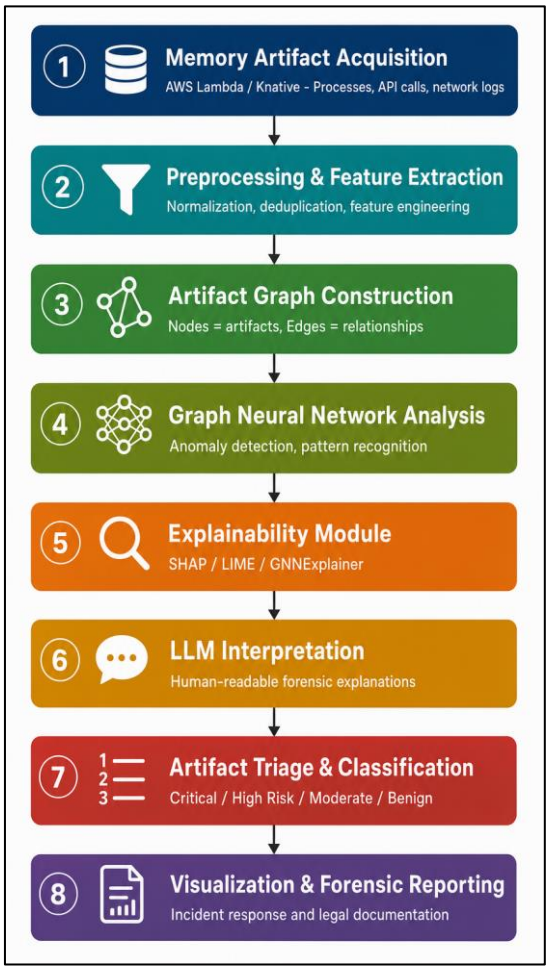

Fig 2 XAI-Based Memory Artifact Triage Framework Pipeline

IV. EXPERIMENTS AND RESULTS

The system's performance is rigorously evaluated using metrics such as accuracy, precision, specificity, recall, F1-score, and confusion matrices to gauge its efficacy in memory artifact triage. The results indicate a robust capability to accurately prioritize evidence, and the model's performance is benchmarked against existing methodologies. The analysis highlights notable trends observed during testing and candidly examines the system's strengths and limitations, offering actionable insights for future enhancements.

Table 1 Outputs of the Memory Artifact Triage System

Artifact Category	Precision	Recall	F1-Score	Accuracy
Critical	0.94	0.92	0.93	0.95
High Risk	0.89	0.91	0.90	0.91
Moderate Risk	0.85	0.84	0.84	0.86
Benign	0.96	0.97	0.96	0.97
Overall	**0.91**	**0.91**	**0.91**	**0.92**

V. DISCUSSION

Deploying XAI for memory artifact triage in serverless cloud forensics necessitates adherence to internationally recognized forensic standards. ISO/IEC 27037 delineates guidelines for the identification, collection, acquisition, and preservation of digital evidence. In serverless environments, forensic validation is exceptionally challenging due to the extreme volatility of memory artifacts, their geographic dispersion, and reliance on third-party providers. Integrating XAI into triage systems fortifies compliance with ISO 27037 by infusing transparency and traceability into evidence analysis. Techniques like SHAP, LIME, and GNNExplainer elucidate why specific artifacts are flagged, fostering evidential accountability and reproducibility—indispensable in legal contexts where courts require a clear understanding of AI-assisted determinations. XAI-enabled systems generate auditable reasoning logs capturing feature importance scores, suspicious relationship mappings, and prediction confidence levels. Nevertheless, full ISO 27037 compliance remains arduous due to opaque infrastructure visibility, necessitating forensic-by-design principles and standardized explainability protocols.

Despite XAI's benefits, scalability remains a profound limitation in ultra-large cloud infrastructures. Platforms like AWS Lambda process millions of invocations rapidly, generating colossal volumes of artifacts and execution logs.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may1257

Applying XAI introduces significant computational overhead, as explainability algorithms require intensive processing to compute feature contributions and decision pathways, potentially inducing latency and escalated operational costs at scale.

Furthermore, integrating XAI in serverless forensics raises critical ethical considerations regarding accountability, fairness, privacy, and trust. Relying on opaque models prevents investigators from justifying conclusions or detecting biases. XAI mitigates this by enabling auditable systems that provide comprehensible rationales, increasing investigator confidence and allowing legal scrutiny. However, biased training data or flawed forensic labeling can yield misleading results. In multi-tenant serverless environments, privacy concerns are amplified as monitoring tools may inadvertently capture sensitive user data. Determining liability for erroneous classifications among the investigator, AI developer, or cloud provider remains complex. Ethical XAI deployment demands a delicate balance between transparency and privacy, automation and human oversight, and efficiency and justice. Future systems must champion responsible AI concepts and sustain human-in-the-loop investigative models.

VI. CONCLUSION AND FUTURE WORK

The application of XAI for memory artifact triage in serverless cloud forensics represents a significant advancement in digital investigations. The proliferation of platforms like AWS Lambda and Knative has introduced novel forensic challenges characterized by ephemerality, scalability, and restricted infrastructure visibility. Embedding XAI into forensic frameworks—utilizing Graph Neural Networks, SHAP, LIME, and Large Language Models—significantly augments investigators' capacity to decipher complex artifact relationships and obtain human-readable rationales for AI decisions. This synergy elevates transparency, trust, and forensic readiness while aiding compliance with ISO/IEC 27037 through auditable reasoning. Despite these strides, hurdles remain regarding ultra-large-scale scalability, real-time evidence retention, computational overhead, and privacy protection. Future research must prioritize lightweight, scalable explainability models, advanced distributed GNN architectures, autonomous forensic orchestration systems, standardized serverless forensics datasets, and blockchain-based immutable logging for chain-of-custody management.

REFERENCES

[1]. F. Idugboe, W. Junior, and V. Castro, "Cloud Forensic Tools and Storage: A Systematic Mapping Study," International Journal of Innovative Science and Research Technology, vol. 3, no. 3, pp. 54–64, 2026. https://doi.org/10.5281/zenodo.19313810
[2]. S. Kavyadharshni, B. Dharciga, and A. S. Vs, "Challenges in Serverless Computing," vol. 03, pp. 1154–1157, 2025.
[3]. S. Kumar and P. Meenalochini, "Explainable AI for Reliable and Transparent Cloud Security Solutions," vol. 2, no. 6, pp. 1–19.
[4]. H. Nyholm, K. Monteith, S. Lyles, M. Gallegos, M. Desantis, J. Donaldson, and C. Taylor, "The Evolution of Volatile Memory Forensics," pp. 556–572, 2022.
[5]. R. D. Syahputri, A. Anggono, and M. Djasuli, "Evolution and Research Opportunities of Digital Forensic Tools: A Bibliometric Analysis," vol. 10, no. 2, pp. 474–485, 2024.
[6]. A. K. Words, D. Learning, and M. Analysis, "A Novel Study on Intelligent Methods and Explainable AI for Dynamic Malware," 2019.
[7]. D. E. Date, "Explainable AI in Malware Analysis and Detection," 2023.
[8]. M. Song, "Explainable AI in Malware Analysis: A Human-Centric Approach," 2023.
[9]. T. Shehzadi, "Serverless Computing Architectures and Applications in AWS," 2023.
[10]. C. James, "Comparative Analysis of Explainable AI Techniques in Malware Detection," 2023.