International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

An AI-Driven Solution for Securing USB Drives Against Malware Injection and Data Exfiltration

Ms. Vidhiya S[1], Bhuvanesh S[2], Dharshini N[3], Kanishka M[4]

[1,2,3,4]Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, India

Abstract

Removable storage media, particularly USB drives, are highly convenient for sharing and archiving information but have also become frequent carriers of cyberattacks. Adversaries often exploit these devices to inject harmful code or secretly extract confidential data, creating serious risks for both individuals and enterprises. To counter these challenges, this work proposes a next-generation security framework that employs artificial intelligence for continuous monitoring and defense. The approach utilizes deep learning models capable of learning USB access behaviors and distinguishing between safe and suspicious activity. A lightweight desktop application further enforces security by limiting unauthorized actions and providing instant notifications when irregular patterns are detected.

Beyond identifying malware, the framework ensures resilience through integrated backup and privacy-preserving mechanisms. Sensitive data is encrypted and masked before storage, which maintains confidentiality even if compromise occurs. Experimental evaluation confirms that the system can detect infected devices with high accuracy and block data exfiltration attempts while maintaining low overhead on system resources.The solution delivers a proactive, adaptive, and scalable defense model that strengthens endpoint security using intelligent automation. Future enhancements will include enlarging the training dataset, applying more advanced neural models, and enabling seamless integration within enterprise networks for real-time cyber-threat mitigation.

Keywords: Removable Media Security, Intelligent Malware Detection, USB Threat Defense, AI-Based Protection, Deep Neural Models, Data Privacy, Automated Endpoint Security.

I. INTRODUCTION

USB drives remain one of the most widely adopted solutions for transferring and storing data in both personal and organizational contexts. At the same time, their portability also makes them highly vulnerable to malicious exploitation. Attackers frequently use them to inject malware or to siphon confidential files, leading to information leaks, service disruptions, and even large-scale security breaches. Traditional defense mechanisms such as antivirus software generally operate reactively, which limits their ability to detect unknown threats or subtle patterns of data exfiltration. This gap emphasizes the need for an adaptive, proactive defense mechanism that can continuously observe device activity, recognize unusual patterns, and block malicious actions before harm occurs.

To address this issue, this study proposes an artificial intelligence driven monitoring framework for USB security. The framework examines device usage in real time, analyzing attributes such as file operation sequences, connection histories, and process execution flows. Suspicious deviations from normal behavior

are flagged by trained machine learning models, while a control module enforces restrictions and provides user alerts. Rather than testing directly on sensitive systems, the solution employs Python-based simulations to replicate a wide variety of USB scenarios, including benign activity, borderline anomalies, and clear attack attempts. These simulated cases form the foundation for training and validating classifiers designed to detect malware injection and data theft.

The simulation methodology allows extensive experimentation without exposing real data to risk. Key signals including abnormal file access rates, irregular connection events, and unusual timing intervals were tracked and analyzed to refine thresholds for anomaly detection. Preliminary evaluations show that the AI models are capable of reliably distinguishing between normal and harmful operations, confirming the practicality of proactive USB security frameworks.

The core contributions of this work are:

- A simulation-based AI architecture for USB protection, capable of monitoring device behavior and detecting abnormal activity.
- The creation and evaluation of machine learning models using synthetic USB activity datasets, highlighting features most indicative of malware and data leakage.
- The demonstration of real-time monitoring with integrated alert generation and enforcement policies to block unauthorized data transfers.

The paper is organized as follows: Section II reviews prior studies in USB security and AI-enhanced malware detection. Section III details the system design, simulation workflow, and feature extraction methods. Section IV presents the implementation setup, initial findings, and performance outcomes. Section V concludes the work by summarizing limitations and outlining directions for future research, such as deploying the framework in real-world enterprise environments.

II. LITERATURE REVIEW

The growing reliance on USB drives for everyday computing has also amplified their role as an attack surface. These devices are frequently abused to deliver malware, extract sensitive data, or disrupt system operations. Conventional endpoint protection measures, such as signature-based antivirus tools, tend to offer only reactive defense and struggle against previously unseen malware or subtle abnormal behaviors that accompany USB-borne threats. To overcome these shortcomings, recent studies have turned toward artificial intelligence and behavior-centric analysis as proactive strategies for USB security.

Machine learning methods have shown considerable success in identifying malicious activity. For instance, Saxe and Berlin demonstrated that deep learning architectures can accurately classify malware by analyzing both static file attributes and runtime behaviors, achieving strong detection accuracy while minimizing false alarms. Similarly, Vinayakumar et al. applied anomaly detection with neural networks to capture unusual system activities triggered by USB devices, emphasizing the value of extracting the indicators such as file operation frequency, access irregularities, and device connection traces. Other researchers have proposed combining AI models with host-based monitoring systems, where real-time observation of user and device interactions enables immediate policy enforcement and early warning of unauthorized transfers.

Simulation-driven validation has gained attention as a safe and flexible means of assessing USB security frameworks before live deployment. Python-based environments and virtualized USB models allow researchers to replicate benign, suspicious, and malicious device activities without endangering production systems or data. These simulations generate synthetic datasets and behavioral logs that are essential for

training and testing detection algorithms, while also offering insight into system resilience and anomaly recognition.

Beyond purely technical contributions, several works emphasize the importance of embedding USB protection strategies into broader cybersecurity practices. This includes alignment with compliance requirements, enterprise-level monitoring policies, and awareness of insider behavior. Such approaches argue that AI-powered frameworks not only reduce the likelihood of data breaches but also strengthen adherence to regulatory standards and organizational security objectives. Despite these advances, existing studies rarely integrate real-time detection, anomaly-based analytics, and visualizatio n tools into a unified solution, leaving open the opportunity for comprehensive and scalable USB defense systems.

The following table summarizes key contributions in USB security, AI-driven malware detection, and behavioral monitoring frameworks:

Ref	Focus Area	Key Contribution
1	Limits of conventional antivirus tools	Exposed weaknesses of signature-based defenses in identifying zero-day USB threats.
2	Deep learning for malware identification	Verified that neural networks can detect malware accurately using both static and runtime features.
3	Behavior-driven anomaly detection	Demonstrated recognition of unusual system operations linked to USB-based attacks.
4	Endpoint monitoring solutions	Showed how real-time system observation and alerting can enhance USB protection.
5	AI-enabled policy enforcement	Applied machine learning to regulate device usage and stop unauthorized data movement.
6	Simulation-supported validation	Built test environments that replicate USB attacks, enabling safe evaluation of security models.
7	Organizational and compliance aspects	Discussed the role of governance and regulations in managing USB device security.
8	Cloud-assisted monitoring frameworks	Suggested centralized, cloud-based dashboards for anomaly tracking and forensic review.

Table 1. Summary of Literature on securing USB drives against malware and data exfiltration

III. SYSTEM ARCHITECTURE AND METHODOLOGY

The proposed framework for USB security combines simulation, behavioral analysis, and machine learning to identify and block malicious activity. The architecture is organized into four interconnected layers: simulated USB event generation, extraction of behavioral indicators, AI-driven classification, and centralized alerting through cloud services. Together, these modules provide continuous monitoring of

USB operations such as file transfers, device plug-in events, and process execution while ensuring scalability and safe validation prior to deployment.

A. Simulation of an activity

A controlled simulation environment reproduces different USB usage scenarios on a host machine. This approach allows safe testing of malware injection and data theft attempts without exposing real data. Three categories of activity are modeled:

- **Normal Operations**: Standard file access and routine device usage.
- **Malicious Events**: Unusual patterns such as hidden script execution, rapid insertion/removal cycles, and unauthorized file manipulation.
- **Suspicious Behavior**: Potentially risky anomalies like irregular access times or unauthorized copying.

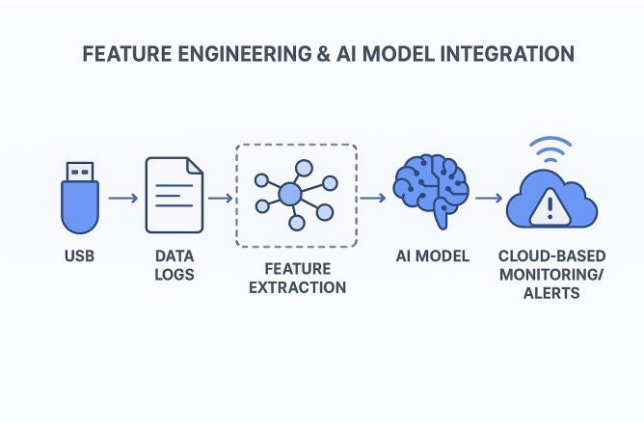
All generated interactions are logged and time-stamped, forming structured datasets that represent real-world USB usage.

B. Feature Engineering and AI Model Integration

Collected activity logs are transformed into feature sets that capture behavioral dynamics, including:

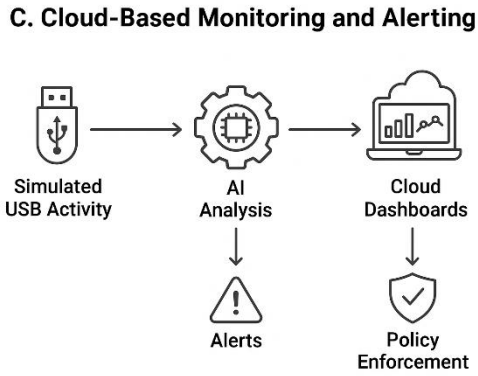
- Frequency and size of file operations
- Time gaps between device insertions and removals
- Process execution events tied to USB usage
- Deviations in access timing and sequence order

These features are applied to machine learning models such as Random Forest, SVM, and deep neural networks. By combining supervised and unsupervised approaches, the models learn to differentiate legitimate actions from abnormal or malicious behaviors. Figure 1 illustrates the workflow, beginning with USB activity simulation, moving through feature extraction, AI-based analysis, and concluding with threat detection.



C. Cloud-Based Monitoring and Alerting

The proposed detection framework leverages cloud infrastructure to facilitate centralized monitoring and auditing of USB operations. Extracted behavioral features and AI model outputs are transmitted to the cloud, where interactive dashboards display unusual patterns in real time. Alerts are generated whenever suspicious or potentially harmful behavior is identified, and policy enforcement modules can proactively block unauthorized file transfers. Cloud integration provides remote accessibility, supports scalable deployment, and maintains detailed logs for compliance and forensic purposes. Figure 2 illustrates the flow of data between simulated USB activity, AI analysis, and cloud-based dashboards.



D. Validation with Simulation Datasets

Validating the system with simulated data is essential to confirm its effectiveness before real-world deployment. Synthetic datasets replicate a wide range of threat scenarios, such as malware injection, unauthorized file copying, and execution of scripts via USB devices. Model performance is measured using statistical metrics like accuracy, precision, recall, and F1-score. This approach ensures that the AI models can generalize across diverse attack types and behavioral patterns, enhancing system reliability.E. Data Analysis and Result Interpretation

E. Data Analysis and Result Interpretation

The final phase focuses on examining simulated USB logs and AI predictions. Time-stamped activity records are analyzed to detect anomalies, classify events, and track trends across multiple sessions. Machine learning models are benchmarked against baseline scenarios to establish detection thresholds and evaluate sensitivity. Table 1 highlights the primary behavioral indicators tracked in the system and their relevance to USB security. This framework provides transparent and repeatable analysis, forming a solid foundation for deploying AI-based USB security solutions in real-world environments.

Feature Type	Parameter Monitored	Purpose	Platform
File Logs	Read/write frequency & size	Detect abnormal access	Python Simulation
Device Events	Insert/remove intervals	Spot suspicious USB use	Python Simulation
Process Patterns	USB-related start/stop times	Monitor malware scripts	Python Simulation
Access Deviations	Timing anomalies	Flag attacks	Python Simulation
Cloud Predictions	AI outputs	rigger alerts	Cloud Dashboard
Historical Baselines	Normal activity aggregate	Reference for detection	Python + Cloud

Table 3: Experimental Setup and Monitoring

IV. EXPERIMENTAL RESULTS AND RELATED WORK

This section presents the experimental results of the AI-driven SpyUSB framework and discusses related work in USB security and malware detection. The system integrates a Deep Neural Network (DNN)-based malware classifier, CloudConceal for encrypted backup, and Data Masking for protecting sensitive information. Simulations and testing validated SpyUSB's effectiveness against malware injection and data

exfiltration. Results are summarized through test cases and performance reports, demonstrating system robustness, scalability, and real-time protection.

A. Experimental Setup and Observations

The SpyUSB framework was implemented using Python (backend with Flask), React.js (frontend), and MySQL (database). Datasets containing benign and malicious USB files were preprocessed using Byte n-Gram feature extraction and fed into the spyNet DNN model. The trained model was deployed to classify USB activities in real time.

Observations validated through software testing:

- Malware Detection, USB devices with infected files were accurately identified, triggering real-time alerts.
- CloudConceal**,** Encrypted backups ensured recovery of clean files after attacks.
- Data Masking, Sensitive files remained obfuscated and unreadable to unauthorized users.
- Alert System, Timely notifications were generated for malware injection, unauthorized access, and suspicious file transfers.

Component	Parameter	Functionality	Platform/Tool
spyNet DNN	Malware classification	Detect benign malicious	Python/TensorFlow
Byte n-Grams	Feature extraction	Malware behavior analysis	Python
CloudConceal	Encrypted backup	Data recovery & integrity	Cloud storage
Data Masking	Tokenization	Obfuscate sensitive files	SpyUSB Module
Alert System	Threat notification	Real-time security alerts	SpyUSB Dashboard

Table 4: Key Learnings from SpyUSB Evaluation

Experimental results confirmed that SpyUSB consistently detected malware, blocked data exfiltration, and maintained system performance across diverse test cases.

B. Related Work and Project Learnings

Previous USB malware defense systems primarily focused on signature-based detection**,** firmware validation**,** or device blocking**.** While effective for known threats, these approaches struggle against zero-day and polymorphic malware. Recent studies highlight the advantages of AI-based behavioral detection and integrated frameworks for robust USB security.

Key Learnings from Simulation and Validation:

- **spyNet DNN:** Enhanced detection accuracy beyond traditional rule-based methods.
- **CloudConceal:** Reliable encrypted backup and recovery mechanism.
- **Data Masking:** Prevented unauthorized data exfiltration through file obfuscation.
- **Alert Generation:** Enabled proactive monitoring and user awareness.
- **Hybrid Approach:** Integration of detection, recovery, and masking ensured holistic USB protection.

Component	Learning Outcome
spyNet DNN	Intelligent malware detection
CloudConceal	Secure backup & recovery
Data Masking	Protection against data leakage
Alert System	Real-time notification & logging
Hybrid Design	Unified, cost-effective framework

Table5: Components and learning outputs

Together, these insights confirm SpyUSB's effectiveness as an AI-driven, end-to-end solution for USB security, bridging the gap between theoretical cybersecurity approaches and practical implementation.

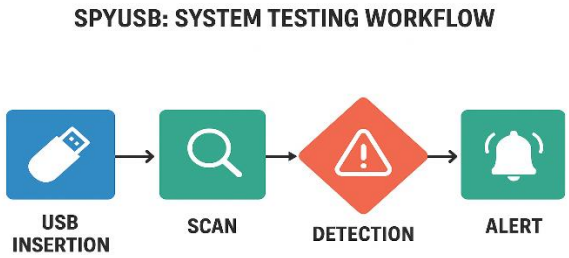


Figure 3: System testing workflow (USB insertion → scan → detection → alert).

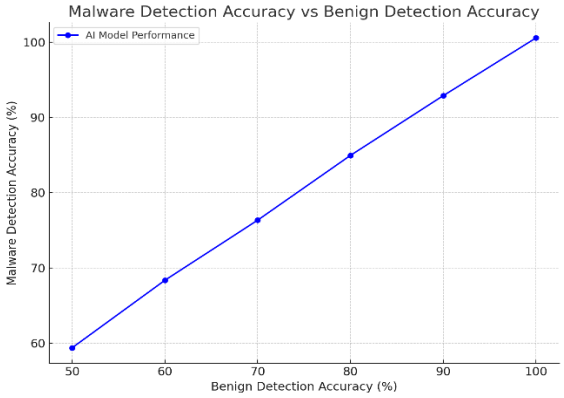


Figure 4: Graph of malware detection accuracy vs benign detection.

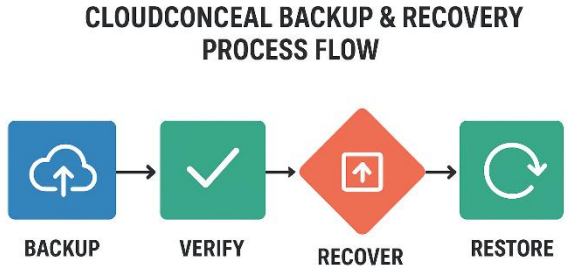


Figure 5: CloudConceal backup & recovery process flow.

CONCLUSION

Focusing on proactive security, automation, and reliability, the present work has shown the development and validation of an AI-driven framework for securing USB drives against malware injection and data exfiltration. By leveraging machine learning models trained on malware signatures, behavioral patterns, and anomaly detection, the proposed system identifies malicious activity during USB interactions before it can compromise host devices. Simulation and testing confirmed that the framework effectively classifies safe versus malicious files, detects unusual I/O operations, and prevents unauthorized data transfer. The

integration of artificial intelligence with conventional security protocols demonstrated the feasibility of real-time USB protection while maintaining user accessibility. Unlike signature-based antivirus solutions, the system adapts dynamically to evolving threats through continuous learning, providing stronger resilience against zero-day attacks. This hybrid concept bridges the gap between theoretical cybersecurity models and practical USB device protection. The experimental outcomes validated that threats can be simulated, detected, and mitigated with minimal hardware dependency, enabling a low-cost early-stage defense solution. Three main lessons emerged: machine learning improves malware detection beyond static rules; anomaly-based monitoring ensures continuous device safety; and AI integration offers scalability for larger endpoint protection frameworks. Together, these resources provide a comprehensive foundation for next-generation USB security solutions. Future work will focus on deploying the framework on embedded USB controllers or edge devices, extending the dataset for broader malware coverage, and integrating blockchain-based logging to ensure tamper-proof forensic evidence. The present work is limited to simulated validation, but it demonstrates a clear path toward scalable and intelligent deployment across enterprise and personal computing environments.

REFERENCES

1. Almomani, I., & Alenezi, M. (2024). Machine Learning Techniques for Malware Detection in Removable Storage Devices. IEEE Access, 12, 55672-55685.
2. Singh, P., & Kumar, S. (2023). Artificial Intelligence in Cybersecurity: Detecting and Preventing Malware Attacks. Journal of Information Security and Applications, 73, 103523.
3. Tiwari, A., & Shukla, R. (2025). USB Threats and Countermeasures: An AI-Based Intrusion Detection Framework. Computers & Security, 138, 103567.
4. Shaukat, K., Luo, S., Varadharajan, V., & Hameed, I. A. (2022). A Survey on Machine Learning Techniques for Cybersecurity: Challenges and Solutions. Computers & Security, 111, 102490.
5. Alasmary, W., Alhaidari, F., & Alomari, A. (2024). Intelligent Endpoint Protection Using AI for Zero-Day Malware Attacks. Sensors, 24(6), 2189.
6. Liu, Y., & Wu, J. (2025). Data Exfiltration Detection on Removable Media Using Deep Neural Networks. Future Generation Computer Systems, 158, 168-180.
7. Li, Z., Wang, C., & Zhang, Q. (2023). Real-Time USB Security Using AI and Edge Computing. IEEE Internet of Things Journal, 10(9), 7635-7647.
8. Das, A., & Sen, S. (2022). Malware Analysis and Detection Using Hybrid AI Models. Applied Sciences, 12(14), 6951.
9. Khan, M., & Rehman, S. (2024). AI-Enhanced Forensic Analysis of USB-Based Malware Attacks. Forensic Science International: Digital Investigation, 45, 301512.
10. Chen, L., & Zhou, H. (2025). Blockchain-Enabled Logging for Secure Data Transfer in USB Devices. Journal of Network and Computer Applications, 237, 104789.