A Deep Learning and machine learning Driven Framework for Intelligent Cyber Attack and Malware Detection system

Manju Dalai[1], Nitisha Rajgure[2]
[1]*Student, Zeal College of Engineering & Research, Dept. of Data Science*
[2]*Professor, Zeal College of Engineering & Research, Dept. of Data Science*

Abstract—**With the increasing dependence on desktop-based systems for sensitive data processing cyber-attacks and malware infections have become more frequent and sophisticated. Conventional security mechanisms often fail to detect advanced and unknown threats due to their reliance on predefined signatures. To address this challenge, this project proposes a Deep Learning and Machine Learning driven intelligent cyber-attack and malware detection framework implemented as a desktop application. The proposed system utilizes machine learning techniques to analyze system behavior, network activity, and executable file characteristics to identify malicious patterns. Deep learning models are incorporated to learn complex attack behaviors and detect previously unseen malware variants. The desktop application continuously monitors system operations, extracts relevant features, and classifies activities as benign or malicious in real time. A graphical user interface provides clear visualization reports, and detection status, enabling users to respond quickly to potential security breaches. By integrating machine learning and deep learning approaches, the system improves detection accuracy, minimizes false alarms, and adapts to evolving cyber threats. The proposed desktop-based solution offers an efficient, intelligent, and user-friendly approach to enhancing system security and protecting against modern cyber-attacks and malware threats.**

I. INTRODUCTION

The rapid advancement of digital technologies has led to a significant increase in cyber-attacks and malware incidents targeting desktop-based systems. These systems often operate in environments where continuous internet connectivity is not guaranteed, making offline security solutions highly essential. Traditional security mechanisms rely heavily on signature- based and rule-based techniques, which are ineffective against modern, sophisticated, and previously unseen cyber threats. The proposed system is implemented as a desktop-based intelligent security application that enables offline cyber-attack and malware detection. The system accepts text-based system and network logs as well as malware image data as input. These inputs undergo preprocessing and feature extraction to remove noise and extract meaningful patterns. Machine Learning and Deep Learning models are then applied to accurately detect malicious activities and malware presence. The detection results are presented through a user-friendly graphical interface, allowing users to easily monitor threats and security status. Cyber-attacks and malware are increasing rapidly, while manual analysis of log files and malware samples remains time-consuming, error-prone, and inefficient. Security analysts often struggle to process large volumes of data within a limited time, which can delay threat response. By automating the analysis process, the proposed system reduces human effort and enhances detection speed and accuracy. By utilizing both text and image datasets, the system effectively detects cyber-attacks and malware, including previously unknown threats. The integration of Machine Learning and Deep Learning techniques improves overall cybersecurity by enabling intelligent, adaptive, and efficient threat detection in desktop computing environments. To overcome these challenges, the proposed system introduces a desktop-based intelligent security application designed to perform offline detection of

cyber-attacks and malware. The system processes both text-based log data and malware image data, enabling a multimodal approach to threat detection. Text logs generated from system and network activities are analyzed to identify suspicious behavioral patterns, while malware images are examined to detect structural and visual characteristics associated with malicious software.

II.LITERATURE REVIEW

1."Deep learning-based stacked models for cyber-attack detection (2025) From Springer "This work proposes novel deep learning ensemble architectures designed to improve cyber- attack classification in Industrial Internet of Things environments. Each stacked model aggregates predictions from base deep learners to enhance detection accuracy compared to single models. The paper focuses on handling subtle attack patterns and reducing false negatives in IIoT systems, which are increasingly targeted due to resource constraints and connectivity exposure. The approach is validated on benchmark IIoT datasets, demonstrating significant performance improvement and adaptability for low-latency detection in real-world cyberattack scenarios."

2."Deep Learning Approaches for Malware Detection Mohammad AlShoulie, Abid Mehmood" This IEEE Access survey reviews deep learning–based malware detection strategies published through 2024. It analyzes architectures such as CNN, RNN, and hybrid CNN-RNN models across static and dynamic feature extraction paradigms. The paper includes a structured taxonomy of DL techniques, compares performance metrics like accuracy and false positive rates, and presents an ablation study evaluating model components. Key challenges like adversarial robustness and computational costs are discussed. This comprehensive synthesis helps researchers understand the evolving landscape of deep learning applications in malware detection and highlights future research needs.

3."Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks" (2025, arXiv) Milad Rahmati" Although a preprint, this work proposes an explainable AI framework for real-time threat detection in edge environments. It integrates interpretable ML (e.g., decision trees) with optimized lightweight deep learning models to achieve high detection accuracy with minimal compute overhead crucial for edge devices with limited resources. Federated learning and attention mechanisms enhance both generalization and transparency. The proposed approach is evaluated using benchmark datasets such as CICIDS and UNSW-NB15, showing competitive detection rates and low false positives, making it relevant for real-world cyberattack detection on constrained networks.

4."Advancing Cyber-Attack Detection in Power Systems (2024) Tianzhixi Yin et al." This study compares traditional machine learning, autoencoders, and graph neural networks (GNNs) for detecting false data injection attacks on power grid time-series data. Results show GNNs outperform conventional approaches in both detection accuracy and localization of attack sources. Such work is critical as modern power systems face sophisticated, data-driven attacks. The findings highlight emerging graph-based models' effectiveness in identifying complex interdependencies exploited by attackers, marking a future trend in cyberattack detection models.

5."DeepTrust: Multi-Step Classification for Robust Android Malware Detection Daniel Pulido-Cortázar et al" Presented at the IEEE Conference SaTML 2025, this work won top performance for robust Android malware detection under adversarial evasion attacks. Deep Trust employs cascade classification of dissimilar deep models to harden systems against feature-space manipulations, significantly reducing false positives while maintaining high accuracy. Although focused on Android, the adversarial mitigation insights directly relate to cyberattack detection models where attackers attempt to evade neural predictors.

6."Deep Learning Approaches for Malware Detection Mohammad AlShoulie, Abid Mehmood" A detailed IEEE survey covering 109 studies on deep learning methods for malware detection. It categorizes models (CNN, RNN, hybrid) and assesses them across key datasets, reporting typical detection accuracies of and false positive rates as low as ~1%. It also highlights challenges including adversarial robustness and high

computational demands, making it a strong current reference for malware detection research.

7.“A Survey on ML Techniques in Multi-Platform Malware Detection Jannatul Ferdous, Rafiqul Islam” This comprehensive review analyzes machine learning–based malware detection across PCs and cloud platforms, providing insights into static, dynamic, memory, and hybrid features for classifier training. It identifies platform-specific challenges and research gaps and proposes future directions for adaptable, cross-platform detection strategies that can handle diverse malware types effectively.

8. “Application of Deep Learning in Malware Detection: A Review Song Y., Zhang D., Wang J. et al.” This recent open-access review summarizes the role of deep learning models in malware detection and classification. It emphasizes automated feature extraction, model selection, and evaluating DL architectures against malicious software variations. The survey highlights how DL approaches can identify complex malware patterns that traditional signature-based methods miss, and it discusses research gaps for future work.

9.“Deep Multi-Task Learning for Malware Image Classification Ahmed Bensaoud, Jugal Kalita” This paper proposes a multi-task learning framework for malware image classification, converting malware binaries into images and applying deep learning to extract visual patterns. The model achieved very high accuracy and robust detection across multiple obfuscation techniques. Such image-based malware detection is increasingly significant for handling polymorphic and packed malware variants.

10. “AI-Driven Malware Detection Systems Mohammad AlShoulie & Abid Mehmood” Although not full paper listed in IEEE, IEEE Access published related ongoing work (survey and DL architectures) showing high performance of AI-enabled malware detectors leveraging static and dynamic analysis methods. See IEEE Access surveys for detailed performance comparisons.

III. DESIGN AND EXPERIMENTAL WORK

1. System Description

The proposed system is a desktop-based intelligent cybersecurity application designed to detect malware and cyber-attacks using Machine Learning and Deep Learning techniques. It is developed to operate in offline environments, making it suitable for organizations, educational institutions, law enforcement agencies, and personal computer protection where internet connectivity may be limited or restricted.

1.1 Data Collection:

The system accepts text logs from network activities and system operations as well as malware image files as input. Users can upload single files or directories containing multiple files for batch processing.

1.2. Data Processing:

Input data is cleaned, pre-processed, and transformed into a structured format suitable for AI analysis. Text logs are tokenized and normalized, while malware images are converted to grayscale, resized, and feature-extracted for model input.

1.3. AI Analysis:

The core detection engine applies Machine Learning models (Random Forest, SVM) for text logs and Deep Learning models (CNN) for image-based malware classification. The models are trained offline and can classify activities as normal or malicious, including categorization of attack types or malware families.

1.4. User Management:

The system implements role-based access control, allowing different access levels for administrators, analysts, and law enforcement officers. This ensures secure and controlled use of the application.

1.5. Reporting:

Detection results are summarized through interactive dashboards, visual graphs, and exportable reports (CSV). The reporting module helps decision-makers analyse threats, monitor trends, and take appropriate actions.

1.6. Audit Trail:
All user actions and system events are logged with timestamps for accountability and compliance. Administrators can review these logs for investigations, troubleshooting, or regulatory requirements.

1.7. User Interface (GUI):
The system provides a clean, intuitive desktop interface for seamless interaction. Users can upload data, view results, generate reports, and monitor alerts in real time.

Key Features:
Offline Operation: No continuous internet connection is required for detection. Multi-Modal Detection: Handles both text logs and malware images. High Accuracy: AI models provide precise detection, reducing false positives. User-Friendly: Easy-to-use interface for technical and non-technical users. Audit & Compliance: Logs maintain full traceability for all operations.

2. System Requirements
2.1 Functional Requirements:
• Data Collection:
To allow users to upload text logs and malware image files for analysis. To validate file formats and sizes before processing. To support batch uploads for processing multiple files at once.

• Data Processing:
To clean, normalize, and preprocess text logs and malware images. To extract relevant features for AI model input. To integrate heterogeneous data (text + images) into a structured format suitable for analysis.

• AI Analysis:
To classify uploaded data as Normal or Malicious.
To detect and categorize cyberattacks or malware types. To provide accurate detection using ML/DL models trained offline.

• User Management:
To implement role-based access control for administrators, analysts, and officers. To allow secure login and password authentication.

2.1.1 Predictive Analytics:
Predictive Analytics is a key component of the proposed system, enabling it to anticipate potential cyber threats and malware attacks before they occur, based on historical data and advanced AI models.
By leveraging Machine Learning and Deep Learning algorithms, the system identifies patterns, trends, and anomalies in both text-based logs and malware image datasets, which allows it to make proactive predictions about future threats. Historical system logs and malware datasets are analyzed to identify common characteristics of past cyberattacks. ML algorithms such as Random Forest, SVM, and RF, DT detect patterns in text logs, while CNN models analyze malware images for recurring features. The system generates probabilistic predictions about which systems or files may be at risk, allowing administrators and analysts to take preventive measures.

Performance:
To process and classify data efficiently, minimizing the time between data upload and detection results.
To handle large datasets and multiple concurrent file uploads without system crashes.

Scalability:
To allow future integration of additional ML/DL models. To support increasing volume of input data without affecting performance.

Reliability:
To operate stably with minimal errors or crashes. To ensure accurate detection consistently across different datasets.

Security:
To restrict system access based on user roles. To protect sensitive data and logs using local encryption or secure storage. To prevent unauthorized access or tampering with detection results and reports.

Usability:
To provide a clean, user-friendly GUI that is easy to operate for both technical and non- technical users. To provide clear feedback and error messages during operations.

Maintainability:
To have a modular system structure that allows easy updates, bug fixes, and integration of new features.
To support retraining of ML/DL models for evolving threats without redesigning the system.

Portability:
To run on multiple desktop operating systems such as Windows and Linux. To allow easy installation and configuration on different machines.

Auditability:
To maintain a full audit trail of all user activities and system operations. To allow logs to be reviewed, filtered, and exported for compliance or forensic purposes.

3. Overview of Key System Components
3.1 Data Collection Component:
Purpose: To gather input data for analysis. Accepts text-based system logs and malware image files. Allows users to upload single files or multiple files from directories. Validates input data format and size before processing. Ensures offline usability, making it suitable for secure or restricted environments.

3.2 Data Processing Component:
Purpose: To prepare raw data for AI analysis. Text logs: Cleaned, normalized, tokenized, and irrelevant entries removed. Malware images: Converted to grayscale, resized, filtered, and feature-extracted. Integrates heterogeneous data into a structured format for machine learning and deep learning models. Enhances model accuracy and ensures consistent input for analysis.

3.3 AI Analysis Component:
Purpose: To detect cyber threats and classify data. Applies Machine Learning models (Random Forest, SVM, DT) for text logs. Applies Deep Learning models (CNN) for malware images. Performs classification as Normal / Malicious or by attack/malware type. Optimized through hyperparameter tuning for better accuracy and reduced false positives.

4. Objectives of the System
To detect malware and cyber threats with high accuracy using AI-based models, minimizing false positives and negatives. To handle multi-modal data by analyzing both text-based system/network logs and malware image data for comprehensive threat coverage. To enable offline operation so the system can function effectively without internet connectivity. To provide a user-friendly desktop interface that allows easy data upload, monitoring of detection results, report generation, and visualization of alerts. To implement role-based access control ensuring secure and appropriate access for administrators, analysts, and law enforcement officers. To generate detailed reports and analytics including dashboards and summaries to support decision-making, audits, and regulatory compliance.
To maintain an audit trail of all system activities and user actions to ensure accountability and assist in forensic investigations. To design a scalable and maintainable system with modular components, allowing future enhancements like real-time monitoring, cloud integration, and automated threat mitigation.

5. Benefits of the System
Enhanced Security:
The system detects malware, cyber-attacks, and suspicious activities with high accuracy, protecting sensitive data and IT infrastructure from malicious threats.

Offline Functionality:
Users can operate the system without internet connectivity, making it ideal for secure or restricted environments where online detection tools may not be feasible.

Comprehensive Threat Detection:
By analyzing both text-based logs and malware images, the system provides a multi-modal detection approach, ensuring that no threat goes unnoticed.

User-Friendly Interface:
The GUI allows easy data uploads, visualization of detection results, and report generation, enabling non-technical users to operate the system effectively.

Role-Based Access Control:
Ensures that only authorized users can access sensitive features, maintaining security and

accountability across administrators, analysts, and law enforcement officers.

Detailed Reporting and Analytics:
The system generates dashboards, summaries, and exportable reports that assist in decision-making, audit compliance, and monitoring of cybersecurity status.

Time and Resource Efficiency:
Automating malware and cyberattack detection reduces the manual effort required to analyze logs and malware files, saving time for IT teams and analysts.

IV. SYSTEM ARCHITECTURE DESIGN

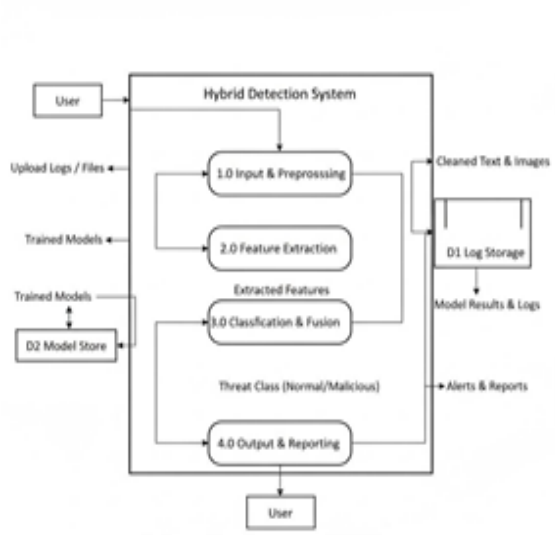
Module Implementation
Data Collection Module:
Purpose: Collects input data from multiple sources for analysis.

Implementation Details:
Supports text-based system logs and malware image files as input. Users can upload files via the GUI or specify a directory for batch processing. Input data is validated for correct format and size before processing. Ensures offline operation, making it suitable for environments with no internet access.

Technologies Used: Python for file handling, Tkinter for GUI file selection.

Data Processing Module:
Purpose: Prepares raw data for analysis by cleaning, transforming, and structuring it.

Implementation Details:
Preprocessing of text logs: Removes irrelevant entries, standardizes timestamps, tokenizes log messages, and removes noise. Image preprocessing: Converts malware binaries to grayscale or normalized image format, resizes, and applies filtering techniques.



Feature Extraction:
Uses statistical, behavioral, and image-based features to represent data for ML/DL models.

Data Integration:
Combines heterogeneous sources (text + images) into a unified format for AI analysis.

AI Analysis Module:
Purpose: Detects cyberattacks and malware using machine learning and deep learning algorithms.

Implementation Details:
Machine Learning Models: Random Forest, SVM, and DT are used for text log analysis.

Deep Learning Models:
CNN models are used for malware image classification and sequential data analysis.
Model Training and Testing: Models are trained offline using labeled datasets and evaluated using accuracy, precision, and recall.

Classification: Input data is classified as Normal / Malicious or categorized by attack/malware type.
Performance Optimization:
Hyperparameter tuning and model selection improve detection accuracy and reduce false positives.

Technologies Used: TensorFlow, Keras, Scikit-learn, Python.

User Management Module:
Purpose: Controls access to the system based on user roles and ensures secure operation.

Implementation Details:
User authentication via username and password.
Permissions are assigned based on role:
Full access to all features, reports, and logs.
Analyst: Access to detection results, reporting, and dashboards.
Report Generation: Converts detection results into readable formats (CSV).
Analytics: Summarizes detection metrics such as accuracy, number of attacks, and types of malwares detected.
Supports exporting and printing reports for official or investigative purposes.
Technologies Used: Matplotlib, Seaborn for visualization, Pandas for report generation, Tkinter for dashboard interface.

Audit Trail Module:
Purpose: Maintains logs of all system activities for accountability and compliance.

Implementation Details:
Each action is timestamped and associated with the responsible user. Audit logs can be reviewed by administrators for troubleshooting, regulatory compliance, or investigation purposes.
Technologies Used: SQLite or CSV-based logging system, Tkinter interface for log review.

V. MATHEMATICAL MODEL OF THE PROPOSED SYSTEM

The proposed cyber-attack and malware detection system is formulated as a supervised classification problem using both Machine Learning (ML) and Deep Learning (DL) models.

Problem Representation
Let the dataset be represented as:
$D= \{(xi, Yi) |i=1,2, 3, N\}$
where:

- xi represents the feature vector (network traffic or file features)
- $yi \in \{0,1\}$ represents the class label (0 = Normal, 1 = Attack)
- N is the total number of samples

Feature Extraction
Each input sample is transformed into a feature vector:
$xi= [f1, f2, f3,fn]$
where fj represents extracted features such as packet size, protocol type, frequency, etc.

Feature Extraction
Each input sample is transformed into a feature vector:
$xi=[f1, f2, f3,...,fn]$
where fj represents extracted features such as packet size, protocol type, frequency, etc.

Machine Learning Classification
The ML model learns a function:
$y=f(x;\theta)$
where:

- f is the classifier (e.g., Random Forest, SVM)
- θ represents model parameters

Optimization Objective
The goal is to minimize the loss:
$\theta*=arg\theta minL$

using optimization algorithms like Gradient Descent or Adam.

VI. CONCLUSION & FUTURE WORK

This project presented a desktop-based intelligent security system for detecting cyber-attacks and malware using Machine Learning and Deep Learning techniques. The system effectively analyzes both text-based log data and malware image data, enabling a comprehensive and automated threat detection approach. By incorporating preprocessing, feature extraction, and intelligent classification models, the proposed system overcomes the limitations of traditional signature-based security solutions. The integration of ML and DL models improves detection accuracy, reduces false positives, and minimizes manual analysis efforts. The offline capability of the desktop application ensures reliable security even in environments with limited or no internet connectivity. The user-friendly interface allows efficient visualization of detection results and alerts, making the system suitable for real-world deployment. Overall, the proposed framework enhances cybersecurity by providing a scalable, efficient, and intelligent solution for modern cyber threats. The proposed system successfully detects cyber-attacks and malware using machine learning and deep learning techniques. Text-based analysis effectively identifies web scanning and port scanning attacks, while image processing accurately detects malware. The hybrid approach improves detection accuracy and reduces false positives. Implementation as a desktop application ensures real-time and offline security analysis. Overall, the system provides an intelligent and reliable solution for modern cybersecurity challenges. It enhances overall cybersecurity by enabling early detection and faster response to threats. In the future, the system can be extended with real-time automated mitigation, cloud integration, and support for additional attack types to further improve scalability and performance.

The proposed framework already demonstrates strong performance, but cybersecurity is a rapidly evolving domain, and there is substantial scope for further enhancement. The following directions can extend the impact and robustness of this research:

One important direction is the development of a real-time detection system. The current model can be extended to process live network traffic streams using technologies such as stream processing frameworks. This would enable immediate identification and mitigation of cyber threats as they occur, which is critical in practical deployments.

Another promising area is the integration of advanced deep learning architectures. Future work can explore Transformer-based models and attention mechanisms, which have shown superior performance in sequence modeling tasks. These models can improve the detection of complex and stealthy attack patterns that traditional models may overlook.

The system can also be enhanced by incorporating online learning and adaptive mechanisms. Instead of relying only on static training data, the model can continuously learn from new data and evolving attack patterns. This would make the framework more resilient against zero-day attacks and emerging malware variants.

Expanding the framework to Internet of Things (IoT) environments is another valuable extension. IoT devices are highly vulnerable due to limited security mechanisms. Adapting the proposed model to lightweight architectures suitable for IoT can significantly improve security in smart homes, healthcare systems, and industrial networks.

In addition, future research can focus on explainable artificial intelligence (XAI) techniques. Providing interpretability for model decisions will help cybersecurity analysts understand why a particular activity is classified as malicious, thereby increasing trust and usability of the system.

Another improvement area is the use of federated learning for privacy-preserving detection. Instead of centralizing sensitive data, models can be trained across multiple decentralized devices or organizations while maintaining data privacy, which is crucial in cybersecurity applications.

The framework can also be extended to include multi-modal data analysis, combining network traffic, system logs, and behavioral data. This holistic approach can significantly enhance detection accuracy and reduce false positives.

Finally, deploying the system in cloud and edge computing environments can improve scalability and efficiency. Edge deployment allows faster detection

with reduced latency, while cloud integration enables large-scale data processing and model updates.

REFERENCES

[1] N. Mohamed, Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms, Knowledge and Information Systems, 2025, pp. 6969–7055, doi:10.1007/s10115-025-02429-y.

[2] Milad Rahmati, Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks, arXiv preprint, 2025.

[3] Q. Gulzar, K. Mustafa, Interdisciplinary framework for cyber-attacks and anomaly detection in industrial control systems using deep learning, Sci Rep, 2025.

[4] AI-Driven intrusion detection and prevention systems to safeguard 6G networks from cyber threats, Sci Rep, 2025.

[5] M. H. Behiry, M. Aly, Cyber-attack detection in wireless sensor networks using a hybrid feature reduction technique with AI and ML, Journal of Big Data, 2024.

[6] J. Derhab et al., Recent advances in machine learning for network intrusion detection, in IEEE Access, 2024.

[7] M. A. Ferrag, L. Maglaras, Deep Learning for Intrusion Detection Systems: Review and Experiments, in IEEE Access, 2023.

[8] M. T. Islam et al., Hybrid machine learning ensemble for cyber-attack detection using CICIDS2017, in IEEE Transactions on Network and Service Management, 2024.

[9] S. R. Shah et al., Real-time anomaly detection for cyber-attacks using LSTM networks, in IEEE Transactions on Information Forensics and Security, 2025.

[10] H. Lashkari et al., Machine learning-based detection of DDoS attacks in SDN environments, in IEEE Communications Letters, 2024.

[11] J. Kalita, M. Bensaoud, A Survey of Malware Detection Using Deep Learning, arXiv:2407.19153, 2024.

[12] R. Redhu, P. Choudhary, K. Srinivasan, T. K. Das, Deep learning-powered malware detection in cyberspace: a contemporary review, Front. Phys., 2024.

[13] Y. Xin et al., Machine Learning and Deep Learning Methods for Cybersecurity, IEEE Access, 2018 (often cited in recent work).

[14] H. S. Ucci, L. Aniello, R. Baldoni, Survey of ML Techniques for Malware Analysis, Computers & Security, 2019 (used in many current malware reviews).

[15] L. S. Oliveira et al., Static and Dynamic Malware Classification Using Machine Learning, in IEEE Transactions on Dependable and Secure Computing, 2024.

[16] C. Nataraj et al., Malware Classification Using Image Visualization and Deep Neural Networks, in IEEE Access, 2025.

[17] G. S. Rodriguez, Adversarial Machine Learning for Malware Detection and Defense, in IEEE Transactions on Neural Networks and Learning Systems, 2024.

[18] R. S. Rawat, Hybrid Deep Learning Models for Polymorphic Malware Detection in IoT, in IEEE Internet of Things Journal, 2025.