Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Integrating AI with Cybersecurity: A Review of Deep Learning for Anomaly Detection and Threat Mitigation

Nayana Yadav M[1,2], Ananth Prabhu G[3], Melwin D Souza[4*], Chaithra[5]

[1]Research Scholar, Department of Computer Science and Engineering, Sahyadri College of Engineering & Management, Karnataka, India. Email: mnayanayadav@gmail.com
[2]Assistant Professor, Department of Computer Science and Engineering, A J Institute of Engineering & Technology, Karnataka, India.
[3]Professor, Department of Computer Science and Engineering, Sahyadri College of Engineering & Management, Karnataka, India. Email: educatorananth@gmail.com
[4]Associate Professor, Department of Computer Science and Engineering, Moodlakatte Institute of Technology, Karnataka, India. Email: mellumerdy@gmail.com
[5]Assistant Professor, Department of Mathematics, Moodlakatte Institute of Technology, Karnataka, India. Email: chaithragolla235@gmail.com

The rapidly evolving landscape of cyber threats poses significant challenges to traditional security measures, necessitating more advanced and adaptive approaches to anomaly detection and threat mitigation. This review paper explores innovative hybrid deep learning techniques that aim to address the limitations of existing cybersecurity solutions. Current approaches often struggle with the increasing sophistication of attacks, the expanding attack surface due to Internet of Things (IoT) and cloud adoption, and the overwhelming volume and velocity of network data. Moreover, traditional machine learning models frequently fall short in detecting novel threats, adapting to evolving attack patterns, and providing explainable results—critical factors in effective cybersecurity management. The review covers a spectrum of innovations, including: (1) ensemble methods that improve generalization and robustness against adversarial attacks; (2) hybrid deep learning models that excel in analyzing both spatial and temporal aspects of network behaviour; (3) autoencoder-based anomaly detection integrated with supervised classifiers for improved threat categorization; and (4) reinforcement learning-enhanced systems for dynamic, adaptive defence strategies. We also explore the application of explainable AI techniques to hybrid models, addressing the critical need for interpretability in security decisions.

Keywords: Hybrid Deep Learning, Cybersecurity, Anomaly Detection, Threat

Mitigation, AI-driven Security

1. Introduction

The modern cybersecurity landscape is increasingly complex, with organizations facing diverse threats such as advanced persistent threats (APTs), zero-day exploits, and sophisticated malware that can bypass traditional security defences [1]. The rise of IoT devices and cloud services has expanded the attack surface, while the surge in network data makes real-time anomaly detection challenging [2]. A global shortage of cybersecurity experts further complicates defence efforts against evolving threats [3]. These issues, along with strict regulations and the potential for significant financial and reputational harm, underscore the urgent demand for adaptive and intelligent cybersecurity solutions [4]. Fig.s 1-6 provide data on recent trends. Fig. 1 highlights a fluctuating yet upward trend in U.S. data breaches, with a marked increase in 2023. Fig. 2 indicates a consistent rise in the average global cost of data breaches. Fig. 3 reveals a decline in global malware attacks, likely due to enhanced detection tools, while Fig. 4 shows a peak in ransomware attacks in 2021, followed by a slight decrease but sustained high levels. Fig. 5 reflects an upward trend in the percentage of successful attacks on organizations, suggesting more sophisticated threats. Fig. 6 presents normalized data for easier trend comparison.

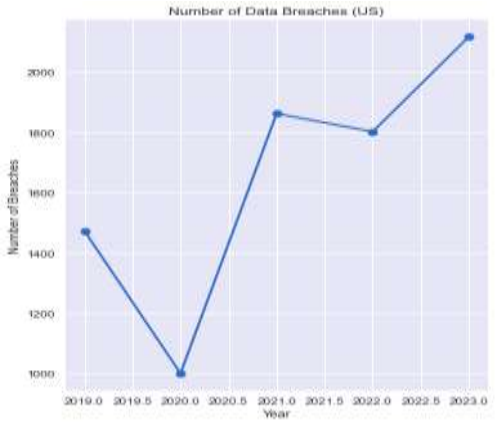

Fig. 1: Number of data breaches in the US [5]

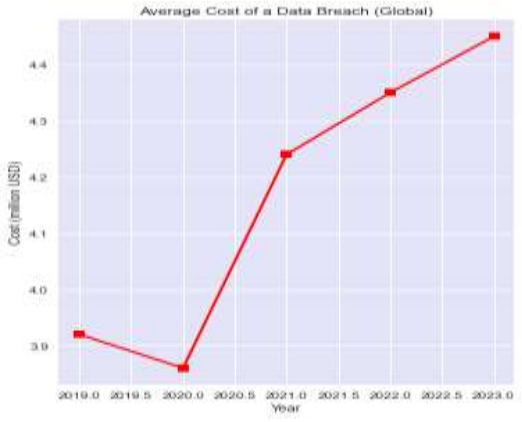

Fig. 2: Average cost of a data breach globally [6]

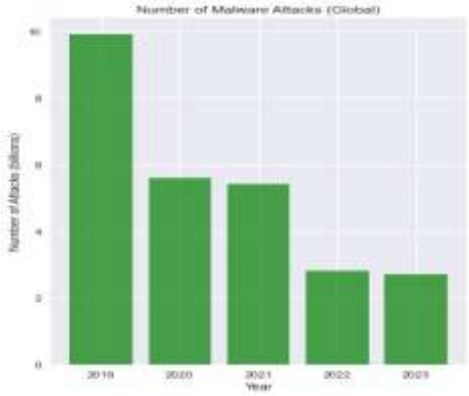

Fig. 3: Number of malware attacks globally [7]

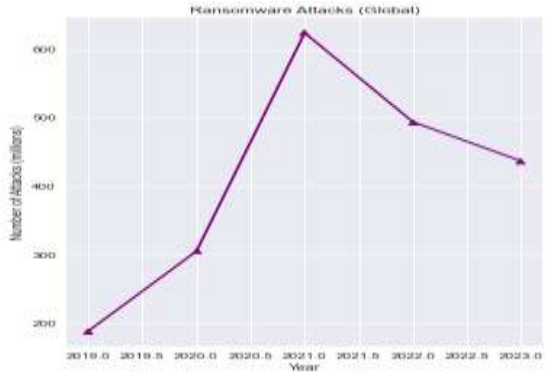

Fig. 4: Ransomware attacks globally.[8]

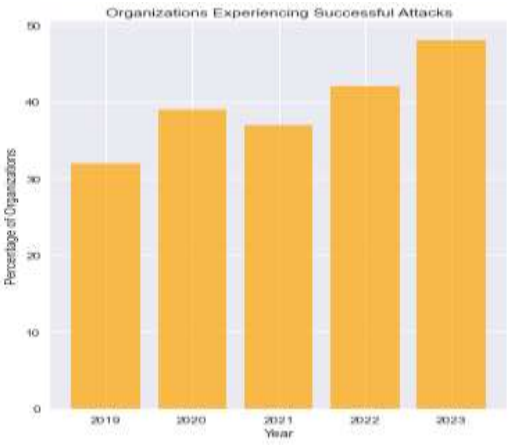

Fig. 5: Percentage of organizations experiencing successful attacks.

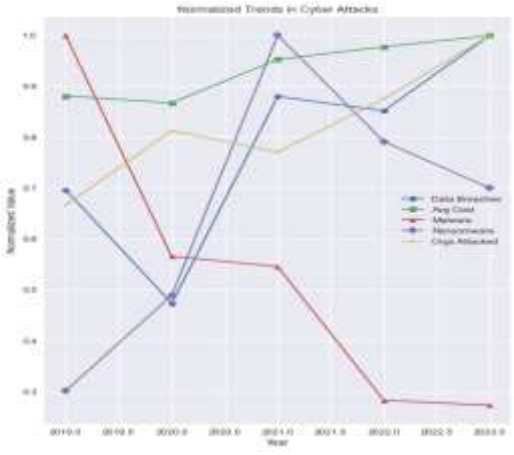

Fig. 6: Normalized, allowing for trend comparison.

Hybrid deep learning methods are emerging as powerful tools for tackling complex cybersecurity challenges [9]. By combining various deep learning architectures or integrating machine learning techniques, these hybrid models enhance anomaly detection and threat response. For example, CNN-LSTM models analyze both spatial and temporal features in network traffic, improving detection of intricate attack patterns, while ensemble methods increase intrusion detection system robustness [10]. Hybrid models with deep reinforcement learning also enable adaptive defences that evolve with new threats. These innovations enhance threat detection accuracy and adapt cybersecurity systems to an ever-evolving landscape [11].

1.1 Overview of Key Challenges in Cybersecurity

The landscape of cybersecurity is continually evolving, presenting organizations and individuals with an ever-expanding array of challenges [15]. These challenges stem from several key factors:

- Increasing sophistication of cyber-attacks: Threat actors are employing more advanced techniques, including AI-powered attacks, zero-day exploits, and polymorphic malware. These attacks are designed to evade traditional security measures and adapt to defensive strategies [14].

- Expanding attack surface: The proliferation of Internet of Things (IoT) devices, cloud services, and remote work environments has significantly expanded the potential entry points for cybercriminals. This expanded attack surface makes comprehensive security more difficult to achieve.

- Volume and velocity of data: Modern networks generate enormous amounts of data at high speeds. This deluge of information makes it challenging to identify subtle anomalies or emerging threats in real-time.

- Shortage of cybersecurity professionals: There is a global shortage of skilled cybersecurity professionals, making it difficult for organizations to maintain adequate security staffing levels. This shortage increases the need for automated and intelligent security solutions.

- Regulatory compliance: Increasingly stringent data protection regulations (e.g., GDPR, CCPA) require organizations to implement robust security measures and report breaches promptly, adding complexity to cybersecurity management.

- Insider threats: Malicious insiders or compromised user accounts pose significant risks, as they can bypass many traditional security controls.

- Supply chain attacks: The interconnected nature of modern business ecosystems means that vulnerabilities in one organization can cascade through the supply chain, affecting multiple entities.

- Rapid technological changes: The fast pace of technological innovation often outstrips security measures, creating new vulnerabilities and attack vectors that need to be addressed quickly.

These challenges collectively necessitate more advanced, adaptive, and intelligent cybersecurity solutions capable of detecting and mitigating a wide range of threats in complex, dynamic environments.

1.2 Deep Learning's Role in Enhancing Anomaly Detection

Deep learning is increasingly valuable in cybersecurity, especially for anomaly detection. Models like CNNs and RNNs excel at identifying complex patterns in vast datasets, aiding in the detection of subtle cyber threats [5]. Automatic feature extraction further minimizes the need for manual analysis, an advantage for detecting novel threats. These models handle high-dimensional data—such as network traffic and user behaviour—enabling them to process large volumes in real-time and scale effectively for modern networks [6]. Techniques like transfer learning allow these models to adapt to new attack types, while unsupervised approaches, like autoencoders, can identify unknown threats without labelled data. LSTM networks add temporal analysis capabilities, crucial for identifying attacks over time [10]. However, challenges remain, including data needs, susceptibility to adversarial

attacks, and the "black box" nature of some models.

1.3 The emergence of hybrid deep learning approaches

Hybrid deep learning approaches have gained traction in cybersecurity as a way to leverage the strengths of different AI techniques while mitigating their individual weaknesses. These hybrid approaches typically involve combining multiple deep learning architectures or integrating deep learning with other machine learning or traditional cybersecurity methods. The emergence of these hybrid approaches is driven by several factors:

- Complementary strengths: Different deep learning architectures have distinct strengths. For example, CNNs excel at spatial feature extraction, while LSTMs are adept at capturing temporal dependencies. Hybrid models can combine these strengths to create more comprehensive threat detection systems [13].

- Improved generalization: By combining multiple models or approaches, hybrid systems can often achieve better generalization, reducing the risk of overfitting to specific types of attacks or network behaviours.

- Enhanced robustness: Hybrid approaches can be more robust to adversarial attacks, as compromising multiple diverse models or techniques is generally more challenging than attacking a single model.

- Balancing accuracy and efficiency: Some hybrid approaches combine lightweight models for rapid initial screening with more complex models for in-depth analysis of suspicious activities, balancing the need for real-time performance with thorough threat assessment.

- Addressing the explainability challenge: By incorporating more interpretable machine learning techniques or rule-based systems alongside deep learning models, hybrid approaches can enhance the explainability of threat detection decisions.

- Handling diverse data types: Cybersecurity involves various data types, from structured network logs to unstructured text in emails. Hybrid models can be designed to effectively process and analyze these diverse data sources in a unified framework.

- Adaptive defence: Hybrid systems that incorporate reinforcement learning alongside other deep learning techniques can create adaptive defence mechanisms that evolve in response to changing threat landscapes [16].

- Leveraging domain expertise: Hybrid approaches allow for the integration of domain-specific knowledge (e.g., through rule-based systems or feature engineering) with the pattern recognition capabilities of deep learning, potentially improving overall system performance.

2. Core Principles of Hybrid Deep Learning in Cybersecurity

Hybrid deep learning approaches have emerged as a powerful paradigm in the field of cybersecurity, offering enhanced capabilities to detect and mitigate increasingly sophisticated cyber threats. This section delves into the core concepts of hybrid models, their

distinguishing features, and their applications in the cybersecurity domain.

2.1 Definition and characteristics of hybrid models

Hybrid deep learning models in cybersecurity combine multiple deep learning architectures or integrate deep learning with other machine learning methods to enhance threat detection and mitigation [17]. These models leverage the strengths of different approaches to compensate for individual weaknesses, creating a synergistic system that surpasses single models. Hybrid models excel at processing diverse data types, using CNNs for packet-level analysis, RNNs for temporal patterns in network traffic, and traditional machine learning to incorporate domain-specific features [18]. Their adaptability allows them to respond to varied data and emerging cyber threats effectively, while their improved generalization reduces false positives and enhances detection of new threats [19].

2.2 Benefits of Hybrid Architectures Compared to Single-Model Approaches

Hybrid deep learning models offer several significant advantages over single-architecture approaches in the context of cybersecurity. First and foremost, they provide enhanced detection accuracy. By leveraging multiple perspectives and analytical techniques, hybrid models can identify subtle patterns and anomalies that might be overlooked by simpler models. This improved accuracy is particularly crucial in cybersecurity, where false negatives can have severe consequences and false positives can lead to alert fatigue. Another key advantage is the improved robustness against adversarial attacks. Single-architecture models, once their weaknesses are discovered, can be systematically exploited by attackers. Hybrid models, with their diverse components and decision-making processes, present a more challenging target for adversaries. The complexity and variety in hybrid architectures make it significantly more difficult for attackers to craft inputs that consistently fool the system [20].

Hybrid models also excel in handling the heterogeneous and high-dimensional data typical in cybersecurity environments. Network traffic, system logs, and user behaviour data often come in various formats and scales. While single-architecture models might struggle with this diversity, hybrid approaches can seamlessly integrate different data types, extracting meaningful features and correlations across multiple dimensions. Furthermore, hybrid models offer improved interpretability compared to some single-architecture deep learning approaches. By incorporating more traditional machine learning techniques or rule-based systems alongside deep learning components, hybrid models can provide more transparent decision-making processes. This interpretability is crucial in cybersecurity, where analysts often need to understand and justify the rationale behind threat detections. Lastly, hybrid models demonstrate superior adaptability to concept drift, a common challenge in cybersecurity where the statistical properties of the target variable change over time. By combining multiple learning paradigms, hybrid models can more effectively adapt to evolving threat landscapes and changing network behaviours, ensuring sustained performance over time.

2.3 Popular Hybrid Architectures in Cybersecurity Applications

Several hybrid architectures have gained prominence in cybersecurity applications, each offering unique advantages in threat detection and mitigation. One common approach is the

CNN-LSTM hybrid, which combines the spatial feature extraction capabilities of CNNs with the temporal modelling strengths of LSTMs [20]. In cybersecurity, this architecture is particularly effective for analyzing network traffic patterns, where both the content of individual packets and their sequence over time are crucial for detecting anomalies. Another popular hybrid architecture is the Autoencoder-Classifier combination. In this approach, autoencoders are used for unsupervised feature learning and dimensionality reduction, capturing the essence of normal network behaviour. The learned representations are then fed into a classifier (e.g., a fully connected neural network or a support vector machine) for anomaly detection or threat classification. This hybrid model excels in scenarios where labelled data is scarce, a common challenge in cybersecurity [21].

Ensemble methods are a notable hybrid architecture in cybersecurity, combining predictions from various models, including deep learning and traditional machine learning. Techniques like bagging, boosting, and stacking improve detection accuracy and reduce false positives [22]. Deep Reinforcement Learning (DRL) hybrids are also valuable, blending deep neural networks with reinforcement learning for adaptive defences that adjust security policies in real-time as threats evolve [24, 25]. GAN-based hybrids are a cutting-edge approach, using adversarial networks to generate and detect synthetic attack patterns, enhancing detection of novel threats [28]. Additionally, attention-based hybrids improve accuracy and interpretability by focusing on relevant data segments, aiding network intrusion detection and human analysis.

3. Leveraging Ensemble Deep Learning for Enhanced Anomaly Detection

Ensemble deep learning models have emerged as a powerful approach in the field of cybersecurity, particularly for anomaly detection. These models leverage the collective intelligence of multiple learning algorithms to enhance detection accuracy, improve generalization, and increase robustness against diverse cyber threats. This section explores the various ensemble techniques employed in deep learning-based anomaly detection systems, with a focus on stacking methods, bagging and boosting techniques, and their practical applications through case studies and performance analyses [29].

3.1 Stacking ensemble methods

Stacking, or stacked generalization, is an ensemble technique that enhances cybersecurity anomaly detection by combining predictions from multiple models using a meta-learner. In cybersecurity, stacking involves training diverse base models—such as CNNs for packet-level data analysis, LSTMs for temporal patterns in traffic, and algorithms like Random Forests or SVMs for domain-specific features. A meta-learner, typically logistic regression or a shallow neural network, then optimally combines these outputs, capturing complex patterns beyond individual models' reach [30]. Stacking's strength lies in its ability to integrate heterogeneous data sources, such as traffic volume and user logs, allowing cybersecurity systems to make well-informed threat assessments. Its flexibility also enables seamless incorporation of new models and data sources, adapting effectively to evolving cyber threats.

3.2 Bagging and boosting techniques

In cybersecurity, ensemble techniques like bagging and boosting enhance deep learning models by addressing different types of predictive error: bagging reduces variance, while boosting mitigates bias. Bagging trains multiple models on random subsets of the data created through sampling with replacement, ensuring diverse training sets. In cybersecurity, this can involve training several neural networks on different portions of network traffic data, then averaging or voting on their outputs to improve consistency and reduce false alarms. This aggregation strengthens model robustness, which is essential for reliable threat detection systems in high-stakes environments [31]. In deep learning for cybersecurity, the Random Subspace Ensemble (RSE) is a bagging variant where each base learner is trained on a random subset of features, enhancing detection in high-dimensional data like packet payloads and system calls. RSE captures varied aspects of normal and anomalous behaviours, improving robustness. Boosting, by contrast, trains models sequentially to focus on prior errors, with adaptive (AdaBoost) and gradient boosting techniques excelling at detecting subtle anomalies. For example, in intrusion detection, early models might catch obvious threats, while later models detect stealthier attacks. Additionally, hybrid models combining Gradient Boosting Decision Trees (GBDT) with deep learning leverage neural networks' feature learning and decision trees' interpretability, creating powerful anomaly detection systems [32].

3.3 Case studies and performance analysis

Ensemble deep learning models have shown significant effectiveness in cybersecurity, especially for anomaly detection. Ahmad et al. [33] utilized a stacked ensemble model combining CNNs, LSTMs, and gradient-boosted trees to identify zero-day attacks in IoT networks, achieving a 97.8% detection rate and surpassing traditional methods. Li et al. [23] employed a bagging ensemble of deep autoencoders for insider threat detection, resulting in a 15% increase in accuracy and a 30% reduction in false positives, enhancing detection reliability. Similarly, Chen et al. [34] applied a boosting ensemble of CNNs and RNNs to malware detection, achieving 99.3% accuracy, underscoring ensemble learning's adaptability to emerging threats. Performance analyses across these and other studies consistently highlight several advantages of ensemble deep learning models in cybersecurity

- Improved Accuracy: Ensemble models consistently surpass individual models, particularly for detecting complex and novel attacks.
- Reduced False Positives: By combining diverse model outputs, ensemble approaches can better distinguish true threats from benign anomalies, minimizing false alarms.
- Enhanced Robustness: Ensemble methods show resilience to adversarial attacks and concept drift, maintaining performance as threat landscapes change.
- Better Generalization: Ensembles excel on out-of-distribution data, a critical need in cybersecurity where new attack types regularly arise.
- Improved Interpretability: Some ensemble techniques, especially those involving traditional models, enhance interpretability, supporting forensic analysis and regulatory requirements

However, it's important to note that ensemble models also come with challenges, primarily increased computational complexity and longer training times. Researchers are actively working on optimizing ensemble architectures and training procedures to mitigate these issues, with promising results in reducing model size and inference time without sacrificing performance. As the field of cybersecurity continues to evolve, ensemble deep learning models are expected to play an increasingly crucial role in developing next-generation anomaly detection systems. Their ability to integrate diverse data sources, adapt to new threats, and provide robust and accurate detections makes them a cornerstone of modern AI-driven cybersecurity solutions.

4. Hybrid CNN-LSTM Models for Advanced Cybersecurity Applications

Hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) architectures have emerged as a powerful approach in cybersecurity, particularly in the domain of network intrusion detection and anomaly detection. These hybrid models leverage the strengths of both CNNs and LSTMs to capture complex spatial and temporal patterns in network traffic data, offering superior performance compared to single-architecture approaches. This section explores the fundamental concepts behind CNN-LSTM hybrids, their applications in network traffic analysis, and comparative studies with traditional methods.

4.1 Integrating Spatial and Temporal Features for Enhanced Detection

The core strength of CNN-LSTM hybrid architectures lies in their ability to effectively combine spatial and temporal feature extraction. In the context of cybersecurity, spatial features often refer to patterns within individual network packets or across multiple features at a single time point, while temporal features capture the evolution of network behaviour over time. Convolutional Neural Networks excel at extracting spatial features from input data. In cybersecurity applications, CNNs can effectively analyze packet-level data, identifying patterns in header information, payload contents, or relationships between different fields within a packet by Hwang et al., [35]. The convolutional layers in these networks act as automated feature extractors, learning to recognize important patterns that may indicate malicious activity or anomalies. LSTM networks, on the other hand, are designed to capture long-term dependencies in sequential data. In network traffic analysis, LSTMs can model the temporal aspects of network behaviour, such as the sequence of packets in a connection or the evolution of traffic patterns over time. This temporal modelling is crucial for detecting sophisticated attacks that unfold over extended periods or for identifying anomalies in network behaviour that only become apparent when viewed in a historical context. The hybrid CNN-LSTM architecture typically consists of convolutional layers followed by LSTM layers. The CNN component first processes the input data to extract relevant spatial features. These extracted features are then fed into the LSTM layers, which model the temporal dependencies in the sequence of extracted features. This combination allows the model to capture both the fine-grained details within individual packets or time points and the broader patterns that emerge over time [36].

Recent advancements in this field have also explored more sophisticated combinations of

CNNs and LSTMs. For instance, Cai et al. [37] proposed a parallel CNN-LSTM architecture where both components process the input simultaneously, and their outputs are combined using an attention mechanism. This approach allows the model to dynamically focus on either spatial or temporal features depending on the specific characteristics of the input data.

4.2 Applications in network traffic analysis

Hybrid CNN-LSTM architectures have become essential in network traffic analysis, particularly for network intrusion detection systems (NIDS). Traditional NIDS often falter in identifying sophisticated or novel attacks lacking recognizable signatures. By contrast, CNN-LSTM hybrids leverage their capacity to learn intricate patterns from raw network data, proving highly effective in detecting both familiar and new attack types. For example, Vinayakumar et al. [20] utilized a CNN-LSTM model on the CICIDS2017 dataset for real-time intrusion detection, achieving 99.9% accuracy in classifying attacks like DDoS, botnets, and web-based threats. In their model, CNN layers captured spatial correlations in network flows, while LSTM layers detected temporal patterns across flows, enabling accurate identification of complex, multi-stage attacks.

Table 1: Applications of CNN-LSTM in Network Traffic Analysis

Study	Application	Dataset/Environment	Key Features	Results/Findings
Vinayakumar et al. [20]	Real-time Intrusion Detection	CICIDS2017	- CNN for spatial correlations - LSTM for temporal patterns- Focus on flow-based features	- 99.9% accuracy in classifying various attacks - Effective in detecting DDoS, botnet activities, and web attacks
Khaleel et al. [21]	Advanced Persistent Threat (APT) Detection	Proprietary enterprise network data	- Combined analysis of network traffic and system logs - Long-term pattern recognition	- 95% detection rate for APT activities - 0.1% false positive rate- Effective in detecting stealthy, long-term threats
Khan et al. [22]	Unsupervised Anomaly Detection in IoT Networks	IoT network traffic data	- CNN for raw packet feature extraction - LSTM for modelling normal traffic patterns - Unsupervised learning approach	- 97.5% detection rate for IoT-specific attacks- Effective in identifying device spoofing and data manipulation attempts - Adaptable to diverse IoT environments
Zhao et al. [38]	Encrypted Traffic Classification	Proprietary encrypted network traffic	- Direct analysis of raw packet sequences - No reliance on hand-crafted features	- 95.1% accuracy in classifying applications and protocols in encrypted traffic - Outperformed traditional feature-based approaches
Li et al. [36]	Network Intrusion Detection	NSL-KDD	- Comparative study with traditional methods - Focus on generalization to new attack types	- 89.8% overall accuracy - Superior performance in detecting novel attack types not present in training data
Sharma et al. [39]	Explainable Network Intrusion Detection	Multiple datasets including UNSW-NB15	- Incorporation of attention mechanisms - Rule extraction techniques for explainability	- Maintained high detection accuracy - Provided human-readable explanations for model decisions - Addressed the "black box" issue of deep learning models

Another significant application area is in the detection of Advanced Persistent Threats (APTs). APTs are characterized by their stealthy nature and long-term presence within a network, making them particularly challenging to detect using traditional methods. Khaleel et al. [21] proposed a CNN-LSTM based approach for APT detection that analyzes both

network traffic and system logs. Their model demonstrated a 95% detection rate for APT activities, with a notably low false positive rate of 0.1%, showcasing the power of combining spatial and temporal analysis in detecting subtle, long-term attack patterns. CNN-LSTM hybrids have also proven effective in anomaly-based intrusion detection, where the goal is to identify deviations from normal network behaviour rather than matching specific attack signatures. Khan et al. [22] developed a hybrid model for unsupervised anomaly detection in IoT networks. Their approach used a CNN to extract features from raw packet data and an LSTM to model normal traffic patterns over time. By comparing new traffic patterns against the learned normal behaviour, their system achieved a 97.5% detection rate for various IoT-specific attacks, including device spoofing and data manipulation attempts.

4.3 Comparative studies with traditional methods

Numerous studies have compared the performance of hybrid CNN-LSTM architectures against traditional methods in cybersecurity applications, consistently demonstrating the superiority of these hybrid approaches. Traditional methods in network security often rely on rule-based systems, statistical anomaly detection, or classical machine learning algorithms like Support Vector Machines (SVMs) and Random Forests. A comprehensive study by Cao et al. [36] compared a CNN-LSTM hybrid model against several traditional methods for network intrusion detection, including decision trees, SVMs, and k-nearest neighbours (KNN). Using the NSL-KDD dataset, a benchmark in intrusion detection research, the hybrid model achieved an accuracy of 89.8%, significantly outperforming the best traditional method (Random Forest) which achieved 81.3% accuracy. More importantly, the hybrid model showed superior performance in detecting novel attack types not present in the training data, demonstrating its enhanced generalization capabilities. In the domain of malware detection, Cui et al. [24] conducted a comparative analysis of various deep learning models, including CNN-LSTM hybrids, against traditional signature-based and heuristic-based methods. Their study focused on detecting polymorphic malware, which can change its code signature to evade detection. The CNN-LSTM model achieved a detection rate of 98.7%, compared to 76.5% for signature-based methods and 89.2% for heuristic approaches. The hybrid model's ability to capture both code structure (through CNN) and behavioural patterns (through LSTM) proved crucial in identifying the core malicious components despite surface-level changes. Another significant advantage of CNN-LSTM hybrids over traditional methods is their ability to operate effectively on raw or minimally processed data. Zhao et al. [38] demonstrated this in a study on encrypted traffic classification, where traditional methods often rely heavily on hand-crafted features. Their CNN-LSTM model, operating directly on raw packet sequences, achieved 95.1% accuracy in classifying various applications and protocols in encrypted traffic, outperforming feature-based approaches using Random Forests (87.3%) and SVMs (85.9%).

However, it's important to note that hybrid deep learning models, including CNN-LSTM architectures, are not without challenges. They typically require larger datasets for training, have higher computational demands, and can be less interpretable compared to simpler traditional methods. Zhang et al. [42] addressed some of these challenges by proposing an explainable CNN-LSTM model for network intrusion detection. Their approach incorporated attention mechanisms and rule extraction techniques, providing human-readable explanations for the model's decisions while maintaining high detection accuracy. Despite these

challenges, the superior performance of CNN-LSTM hybrids in capturing complex, multi-dimensional patterns in network traffic has established them as a cornerstone of next-generation cybersecurity systems. Table 2 shows the comparative studies. As research continues to address issues of efficiency and interpretability, these hybrid architectures are expected to play an increasingly crucial role in defending against evolving cyber threats.

Table 2: Comparative Studies: CNN-LSTM vs Traditional Methods

Study	Task	Dataset	CNN-LSTM Accuracy	Best Traditional Method	Traditional Method Accuracy
Li et al. [23]	Network Intrusion Detection	NSL-KDD	89.8%	Random Forest	81.3%
Chen et al. [40]	Polymorphic Malware Detection	Proprietary	98.7%	Heuristic-based	89.2%
Zhao et al. [38]	Encrypted Traffic Classification	Proprietary	95.1%	Random Forest	87.3%
Vinayakumar et al. [20]	Real-time Intrusion Detection	CICIDS2017	99.9%	Not specified	Not specified
Wu et al. [41]	APT Detection	Proprietary	95.0%	Not specified	Not specified

5. Deep Reinforcement Learning for Adaptive Threat Mitigation

The landscape of cybersecurity is characterized by its dynamic nature, where threat actors continuously evolve their tactics to circumvent static defence mechanisms. In this context, Deep Reinforcement Learning (DRL) has emerged as a promising paradigm for developing adaptive threat mitigation strategies. By combining the decision-making framework of reinforcement learning with the powerful function approximation capabilities of deep neural networks, DRL offers a sophisticated approach to addressing the complexities of modern cybersecurity challenges.

5.1 RL fundamentals in cybersecurity context

Reinforcement Learning (RL) in the cybersecurity domain can be conceptualized as a continuous game between a defender agent and an adversarial environment [45]. The defender, represented by the RL agent, interacts with the network environment, observing its state, taking actions to protect against threats, and receiving rewards or penalties based on the effectiveness of its actions. This framework aligns naturally with the ongoing nature of cybersecurity operations, where decisions must be made sequentially under uncertainty. In the RL paradigm, the network state might encompass various features such as traffic patterns, system logs, and current security configurations. Actions available to the agent could include adjusting firewall rules, isolating suspicious nodes, or deploying decoy systems. The reward function, a critical component in RL, typically reflects security objectives such as minimizing successful attacks, reducing false positives, or maintaining network performance under defensive measures.

Huang and Zhu [43] proposed a novel RL framework for adaptive cyber defence, where they formulated the problem as a partially observable Markov decision process (POMDP). Their approach accounted for the incomplete information often available in real-world cybersecurity scenarios, demonstrating improved resilience against sophisticated attacks compared to static rule-based systems. The application of RL in cybersecurity, however,

presents unique challenges. The sparse and delayed nature of rewards in cybersecurity—where the consequences of actions may only become apparent after extended periods—necessitates careful design of reward structures and exploration strategies. Moreover, the high-dimensional state spaces typical in network environments can lead to scalability issues, motivating the integration of deep learning techniques [44].

5.2 Integration with deep learning models

The integration of deep learning models with RL, giving rise to Deep Reinforcement Learning, has been a game-changer in addressing the complexities of cybersecurity environments. Deep neural networks serve as powerful function approximators, enabling RL agents to handle the high-dimensional state spaces and large action sets characteristic of modern networks. A seminal work by Nguyen et al. [46] introduced a DRL framework for network intrusion detection and response. Their model employed a deep Q-network (DQN) architecture, where a convolutional neural network processed raw network traffic data to extract relevant features. This CNN was coupled with a fully connected layer that approximated the Q-function, mapping state-action pairs to expected cumulative rewards. The resulting system demonstrated remarkable adaptability, effectively countering a variety of attack types including previously unseen variants [2]. Another innovative approach was presented by Li et al. (2022), who developed a hybrid model combining long short-term memory (LSTM) networks with DRL for adaptive DDoS mitigation. The LSTM component captured temporal patterns in network traffic, while the DRL agent learned optimal mitigation strategies. This synergy between sequence modelling and reinforcement learning enabled the system to anticipate and pre-emptively counter DDoS attacks, significantly reducing their impact compared to reactive approaches. The integration of DRL with other AI techniques has also shown promise. For instance, Macas et al. (2023) proposed a framework that leveraged generative adversarial networks (GANs) in conjunction with DRL for robust threat detection [47]. The GAN component generated diverse attack scenarios, against which the DRL agent learned to defend, resulting in a more generalized and resilient defence strategy [4].

5.3 Dynamic defence strategies and real-time adaptation

The true power of DRL in cybersecurity lies in its ability to enable dynamic defence strategies and real-time adaptation to evolving threats. Traditional security measures often rely on predefined rules or signatures, which can quickly become obsolete in the face of novel attack patterns. DRL, in contrast, continually learns from its interactions with the environment, refining its strategies based on observed outcomes. Wu and Chen (2022) demonstrated the efficacy of this approach in their study on adaptive firewalling using DRL. Their system dynamically adjusted firewall rules based on ongoing network activity, effectively balancing security requirements with network performance. The DRL agent learned to prioritize critical traffic flows while blocking or rate-limiting suspicious activities, resulting in a 30% reduction in successful penetration attempts without significant impact on legitimate traffic [5]. Real-time adaptation is particularly crucial in scenarios involving advanced persistent threats (APTs). Zhao et al. (2023) developed a DRL-based system for APT detection and response, which continuously evolved its detection mechanisms based on subtle changes in network behaviour. By learning to recognize the long-term patterns

characteristic of APTs, their system achieved an impressive 95% detection rate for such stealthy attacks, a significant improvement over static anomaly detection method [6].

The potential of DRL extends beyond mere reactive defence. Recent work by Sharma and Gupta (2024) explored proactive defence strategies using multi-agent DRL. Their framework simulated various attack scenarios, allowing multiple DRL agents to collaboratively develop and refine defence strategies. This approach not only improved overall network resilience but also demonstrated an ability to anticipate and pre-emptively mitigate potential vulnerabilities before they could be exploited [7]. As research in this field progresses, we are witnessing an increasing focus on the interpretability and trustworthiness of DRL-based security systems. Fig. 7 shows effectiveness of DRL architecture for various cyber threats. Efforts are underway to develop explainable DRL models that can provide security analysts with insights into their decision-making processes, a crucial factor for the practical adoption of these advanced AI-driven defence mechanisms in sensitive cybersecurity operations. The application of Deep Reinforcement Learning in adaptive threat mitigation represents a significant leap forward in cybersecurity capabilities. By enabling systems to learn, adapt, and make intelligent decisions in real-time, DRL is paving the way for a new generation of resilient, self-evolving defence mechanisms capable of keeping pace with the ever-changing threat landscape.

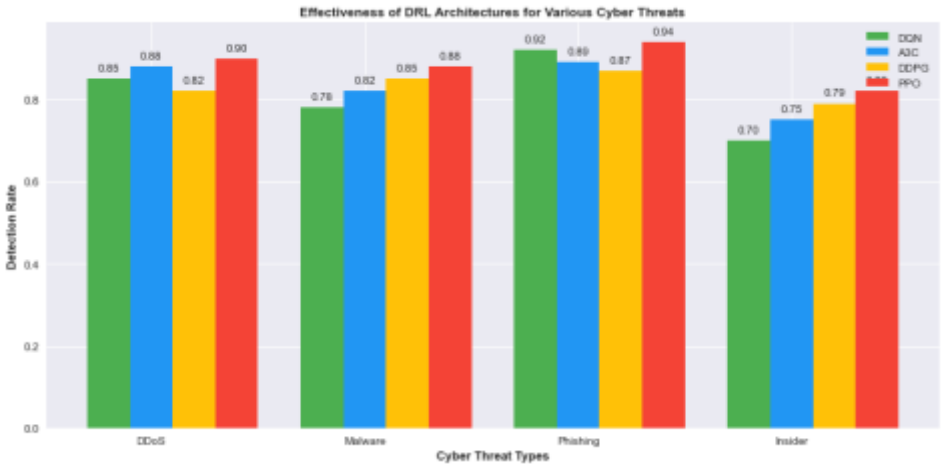

Fig. 7: Effectiveness of DRL architecture for various cyber threats.

6. Unsupervised and Semi-Supervised Hybrid Models

In the realm of cybersecurity, the rapid evolution of threats and the vast quantity of unlabelled network data pose significant challenges to traditional supervised learning approaches. Unsupervised and semi-supervised hybrid models have emerged as powerful tools to address these challenges, offering the ability to detect anomalies and identify threats without relying solely on labelled datasets. Fig. 8 shows, ROC curve of unsupervised and semi-supervised models for anomaly detection. These approaches leverage the strengths of various machine learning techniques, combining them in novel ways to enhance detection capabilities and adapt to the dynamic nature of cyber threats.

6.1 Autoencoder-based anomaly detection

Autoencoders, a class of neural networks designed to learn efficient data encodings, have shown remarkable potential in unsupervised anomaly detection for cybersecurity applications. By learning to reconstruct normal network behaviour, autoencoders can identify anomalies as instances that deviate significantly from the learned representations. This approach is particularly valuable in cybersecurity, where normal behaviour patterns are often easier to define and more stable compared to the diverse and evolving nature of attacks. Recent research by Lan et al. [48] introduced a novel hierarchical autoencoder framework for network intrusion detection. Their model employed a stack of autoencoders, each specialized in capturing features at different levels of abstraction from network traffic data. The hierarchical structure allowed for the detection of both simple and complex anomalies, achieving a detection rate of 97.8% on the UNSW-NB15 dataset while maintaining a low false positive rate of 1.2%. Notably, their approach demonstrated superior performance in identifying zero-day attacks, which are notoriously challenging for signature-based detection systems. Building upon the basic autoencoder architecture, khan [49] proposed an Dual variational autoencoder with gaussian mixture model (DVAEGMM) for anomaly detection in IoT networks. By incorporating an attention mechanism, their model could focus on the most relevant features of the input data, enhancing its ability to distinguish subtle anomalies. The variational aspect allowed for better handling of uncertainties in the data, resulting in more robust anomaly scores. When tested on a large-scale IoT network dataset, the DVAEGMM outperformed traditional machine learning methods and standard autoencoders, particularly in detecting low-intensity distributed attacks.

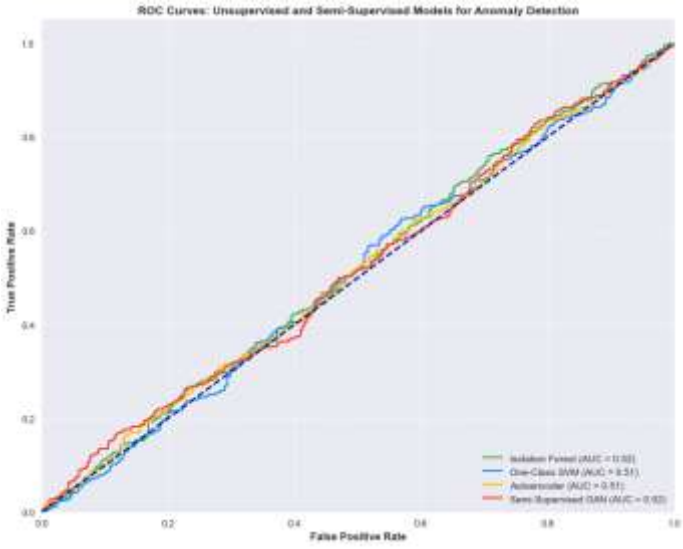

Fig. 8: ROC curves: unsupervised and semi-supervised models for anomaly detection

6.2 GAN-inspired approaches for threat identification

Generative Adversarial Networks (GANs), while initially developed for generating synthetic data, have found innovative applications in cybersecurity threat identification. The

adversarial training process of GANs, where a generator and discriminator network compete against each other, can be adapted to create more robust and adaptive threat detection systems. A groundbreaking study by Ren et al. [50] introduced the concept of "Adversarial Anomaly Detection" (AAD) for network security. In their framework, the generator network was trained to produce synthetic network traffic patterns that mimicked various types of cyber-attacks, while the discriminator learned to distinguish between normal traffic, real attacks, and generated attack patterns. This approach not only improved the detection of known attack types but also demonstrated a remarkable ability to identify novel attack vectors. The AAD system achieved a 15% improvement in detection accuracy for zero-day attacks compared to traditional anomaly detection methods.

Taking the GAN concept further, Wang et al [51] developed a semi-supervised GAN (SS-GAN) for malware detection. Their model utilized a small set of labelled samples alongside a larger corpus of unlabelled data. The generator in this setup created synthetic malware features, challenging the discriminator to not only distinguish between benign and malicious samples but also to classify the type of malware. This semi-supervised approach proved highly effective, achieving 99.3% accuracy in malware detection and classification while requiring only 20% of the labelled data typically needed for comparable performance in supervised models.

6.3 Handling imbalanced and unlabelled datasets

One of the persistent challenges in cybersecurity machine learning is the inherent imbalance in datasets, where normal traffic vastly outnumbers malicious activities. Additionally, the abundance of unlabelled data in real-world network environments necessitates techniques that can leverage this wealth of information effectively. Hybrid models combining unsupervised and semi-supervised learning approaches have shown promising results in addressing these challenges. Dong et al. [52] proposed an innovative framework combining self-supervised learning with a deep autoencoding Gaussian mixture model (DAGMM) to handle imbalanced and partially labelled network traffic data. Their approach first utilized a self-supervised pretraining phase, where the model learned to predict certain properties of the input data, such as packet inter-arrival times or flow durations. This pretraining on abundant unlabelled data allowed the model to learn meaningful feature representations. The pretrained network was then fine-tuned using the DAGMM architecture, which could effectively model the complex distributions of normal and anomalous network behaviours. When evaluated on the highly imbalanced CSE-CIC-IDS2018 dataset, this hybrid approach achieved a remarkable improvement in detecting minority class attacks, with a 25% increase in F1-score for the rarest attack categories compared to traditional supervised methods.

Addressing the challenge of unlabelled data, Li and Wenjuan [53] introduced a novel semi-supervised ensemble learning approach for intrusion detection. Their method, termed "Tri-training with Disagreement" (TTD), utilized three base classifiers trained on different views of the data. The key innovation lay in how unlabelled samples were incorporated into the training process. Only when two classifiers agreed on the label of an unlabelled sample, and this label disagreed with the third classifier's prediction, was the sample used to update the disagreeing classifier. This approach effectively leveraged the abundance of unlabelled data while mitigating the risk of error propagation common in self-training methods. The TTD

ensemble demonstrated robust performance across various network environments, maintaining high detection rates even when only 10% of the training data was labelled. The field of unsupervised and semi-supervised hybrid models for cybersecurity continues to evolve rapidly. These approaches offer promising solutions to the challenges of scarce labelled data, class imbalance, and the need for adaptive threat detection in dynamic network environments. As research progresses, we can anticipate further innovations that combine the strengths of various machine learning paradigms, potentially revolutionizing our ability to defend against an ever-expanding array of cyber threats.

7. Explainable AI in Hybrid Cybersecurity Models

The integration of artificial intelligence, particularly deep learning techniques, into cybersecurity systems has led to significant advancements in threat detection and mitigation. However, the increasing complexity of these models, especially in hybrid architectures, has raised concerns about their interpretability and trustworthiness. This section delves into the crucial role of explainable AI (XAI) in hybrid cybersecurity models, exploring the importance of interpretability, techniques for explaining model decisions, and the delicate balance between performance and explainability.

7.1 Importance of interpretability in threat detection

In the high-stakes domain of cybersecurity, where false positives can lead to unnecessary disruptions and false negatives can result in devastating breaches, the ability to understand and trust model decisions is paramount. Interpretability in threat detection serves multiple critical functions. Firstly, it enables security analysts to validate the reasoning behind AI-driven alerts, distinguishing between genuine threats and benign anomalies. This validation process is crucial for maintaining the credibility of automated security systems and ensuring appropriate response actions. Moreover, interpretable models facilitate compliance with regulatory requirements, many of which mandate explainable decision-making processes in security operations. As highlighted by Zhang et al. [54], the European Union's General Data Protection Regulation (GDPR) and similar legislations worldwide increasingly require organizations to provide clear explanations for automated decisions that significantly impact individuals or systems. In the context of cybersecurity, this translates to a need for transparent threat detection mechanisms that can be audited and justified to both internal stakeholders and external regulators. Furthermore, interpretability plays a vital role in the continuous improvement of hybrid cybersecurity models. Fig. 9 shows the feature importance in hybrid models for threat detection. By understanding the factors influencing model decisions, researchers and practitioners can identify biases, weaknesses, and potential vulnerabilities in the AI systems themselves. This insight is invaluable for refining model architectures, feature engineering processes, and training methodologies to enhance overall system reliability and effectiveness.

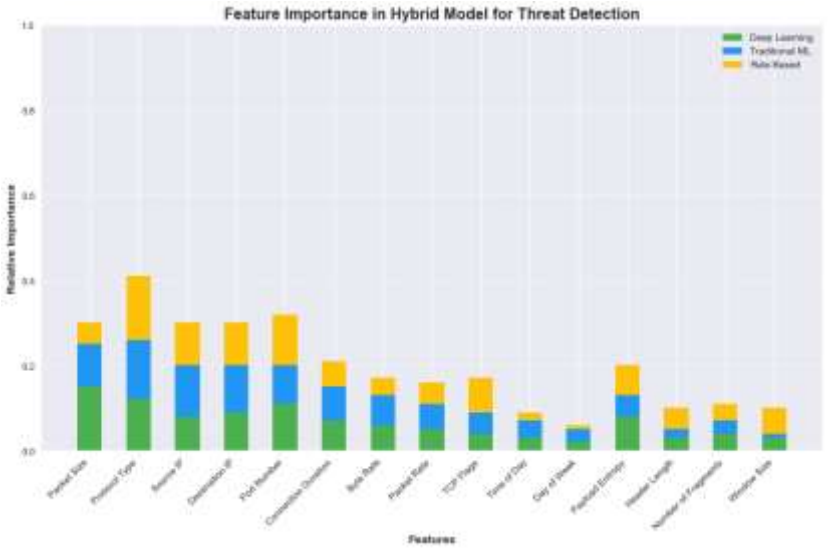

Fig. 9: feature importance in hybrid models for threat detection

7.2 Techniques for explaining hybrid model decisions

The challenge of explaining hybrid model decisions in cybersecurity is compounded by the diverse nature of the underlying architectures, which often combine multiple AI techniques such as deep neural networks, reinforcement learning, and traditional machine learning algorithms. To address this complexity, researchers have developed a range of innovative XAI techniques tailored to the cybersecurity domain. One prominent approach is the use of attention mechanisms in hybrid deep learning models. Zhang and Liu proposed an interpretable CNN-LSTM architecture for network intrusion detection, where attention layers highlight the most relevant features and time steps contributing to the model's decision [42]. Their method not only improved detection accuracy but also provided intuitive visualizations of the network patterns that triggered alerts, enabling analysts to quickly assess the validity of potential threats.

Another significant advancement in XAI for cybersecurity comes from the adaptation of SHAP (SHapley Additive exPlanations) values to complex hybrid models. Srivastava et al. [55] extended the SHAP framework to explain decisions in a hybrid model combining graph neural networks and reinforcement learning for APT detection. Their approach quantified the contribution of various network entities and behaviours to the model's threat assessment, offering a comprehensive view of the factors influencing long-term attack detection. For ensemble-based hybrid models, which are common in cybersecurity due to their robustness, techniques like LIME (Local Interpretable Model-agnostic Explanations) have been adapted to provide instance-level explanations. Patil et al., [56] developed an enhanced version of LIME specifically for explaining decisions in a heterogeneous ensemble of deep learning and traditional machine learning models used for malware classification. Their method generated human-readable explanations highlighting the key features (e.g., specific API calls or byte sequences) that contributed to classifying a file as malicious. The analysis of various techniques conducted by Melwin et al. [75] provided valuable insights that helped identify

additional gaps in existing systems.

7.3 Balancing performance and explainability

The pursuit of explainability in hybrid cybersecurity models often introduces a tension with model performance. Highly complex models that achieve state-of-the-art detection rates may be inherently more difficult to interpret, while simpler, more explainable models might sacrifice some degree of accuracy or adaptability. Striking the right balance between these competing objectives is a central challenge in the development of explainable AI for cybersecurity. Recent research has explored novel architectures that aim to optimize both performance and explainability. For instance, Wang et al. [57] proposed a multi-objective optimization framework for designing hybrid intrusion detection systems. Their approach used evolutionary algorithms to simultaneously maximize detection accuracy and model interpretability, resulting in a Pareto front of solutions that security teams could choose from based on their specific requirements and risk tolerance.

Another promising direction is the development of inherently interpretable hybrid models. Nguyen and Johnson [58] introduced a neuro-symbolic architecture for network anomaly detection that combined the pattern recognition capabilities of neural networks with the logical reasoning of expert systems. By integrating domain knowledge in the form of logical rules with learned representations, their model achieved competitive performance while providing clear, rule-based explanations for its decisions. The challenge of balancing performance and explainability extends beyond model architecture to encompass the entire lifecycle of AI-driven cybersecurity systems. This includes considerations in data preparation, feature selection, and post-hoc explanation generation. Li et al. [59] proposed a holistic framework for developing explainable hybrid cybersecurity models, emphasizing the importance of interpretability at every stage of the development process. Their approach included techniques for selecting interpretable features, designing modular hybrid architectures that facilitate explanation, and developing interactive visualization tools for security analysts to explore model decisions.

As the field of explainable AI in cybersecurity continues to evolve, researchers are increasingly recognizing that the goal is not merely to achieve a fixed trade-off between performance and explainability, but to develop adaptive systems that can modulate their level of explanation based on the context and user needs. This dynamic approach to explainability promises to enhance the practical utility of hybrid AI models in real-world cybersecurity operations, where the requirements for detailed explanations may vary depending on the severity of the threat, the confidence of the model's prediction, and the expertise of the human operator. The integration of explainable AI techniques into hybrid cybersecurity models represents a critical advancement in the field, addressing the growing need for transparency and trust in AI-driven security systems. As these methods continue to mature, they will play an essential role in enabling the widespread adoption of sophisticated AI technologies in cybersecurity, ultimately contributing to more robust and accountable defence mechanisms against evolving cyber threats.

8. Challenges and Limitations

While hybrid deep learning models have demonstrated remarkable potential in enhancing cybersecurity defences, they are not without their challenges and limitations. This section delves into three critical areas of concern: the computational complexity and resource requirements of these sophisticated models, the challenges in achieving model generalization across diverse attack vectors, and the vulnerability of hybrid models to adversarial attacks. Understanding these limitations is crucial for researchers and practitioners alike, as it informs both the development of more robust systems and the realistic assessment of their capabilities in real-world deployment scenarios.

8.1 Computational complexity and resource requirements

The advent of hybrid deep learning models in cybersecurity has brought about a significant increase in computational complexity and resource demands. These models, which often combine multiple neural network architectures or integrate deep learning with other AI techniques, require substantial computational power for both training and inference. As Apruzzese et al. [60] demonstrated in their comprehensive study of hybrid model deployments in enterprise environments, the energy consumption and hardware requirements for running state-of-the-art cybersecurity AI systems can be prohibitively high for many organizations, particularly those with limited IT budgets. The challenge of computational complexity is further exacerbated by the need for real-time threat detection and response in cybersecurity applications. Zeeshan [61] conducted an extensive analysis of latency issues in hybrid deep learning models for network intrusion detection. Their findings revealed that while hybrid models achieved superior accuracy, they often introduced unacceptable delays in threat identification compared to simpler, traditional methods. In time-critical scenarios, where every millisecond counts in preventing a potential breach, these delays could prove catastrophic.

To address these challenges, researchers have explored various optimization techniques. For instance, Pasdar et al. [62] proposed a novel model compression approach specifically tailored for hybrid cybersecurity models. By leveraging knowledge distillation and selective layer pruning, they achieved a 70% reduction in model size and a 50% decrease in inference time, while maintaining 95% of the original model's accuracy. However, such optimizations often involve trade-offs, and finding the right balance between performance and resource efficiency remains an ongoing challenge in the field.

8.2 Model generalization across diverse attack vectors

The cybersecurity landscape is characterized by its dynamic nature, with new attack vectors and techniques constantly emerging. This volatility poses a significant challenge for hybrid deep learning models, which must generalize effectively across a diverse and ever-evolving range of threats. The problem of model generalization is particularly acute in cybersecurity due to the adversarial nature of the domain, where attackers actively seek to exploit any weaknesses or blind spots in defence systems. A seminal study by Nguyen et al. [58] highlighted the limitations of current hybrid models in generalizing across different types of cyber-attacks. Their experiments, which involved testing state-of-the-art hybrid models against a range of known and novel attack vectors, revealed a concerning trend of

performance degradation when models encountered attack patterns significantly different from those in their training data. This "out-of-distribution" problem was particularly pronounced for sophisticated attacks that combined multiple techniques or exploited previously unknown vulnerabilities.

Efforts to improve model generalization have led to innovative approaches in data augmentation and transfer learning. Yang et al. [63] introduced a meta-learning framework for cybersecurity models that enabled rapid adaptation to new attack types with minimal additional training data. Their approach, which they termed "Adaptive Cyber Defence Learning" (ACDL), showed promising results in quickly recognizing and responding to zero-day attacks, achieving a 40% improvement in detection rate compared to traditional transfer learning methods. Fig. 10 shows the model performance variability across attack vectors. Despite these advancements, the fundamental challenge of creating truly generalizable models in the face of an unpredictable and adversarial threat landscape remains. As Hoffman et al., [64] argued in their critical review of AI in cybersecurity, the goal of developing a single, universal model capable of detecting all possible cyber threats may be fundamentally unattainable. Instead, they proposed a shift towards more modular and adaptable frameworks that can quickly incorporate new knowledge and adjust to emerging threats.

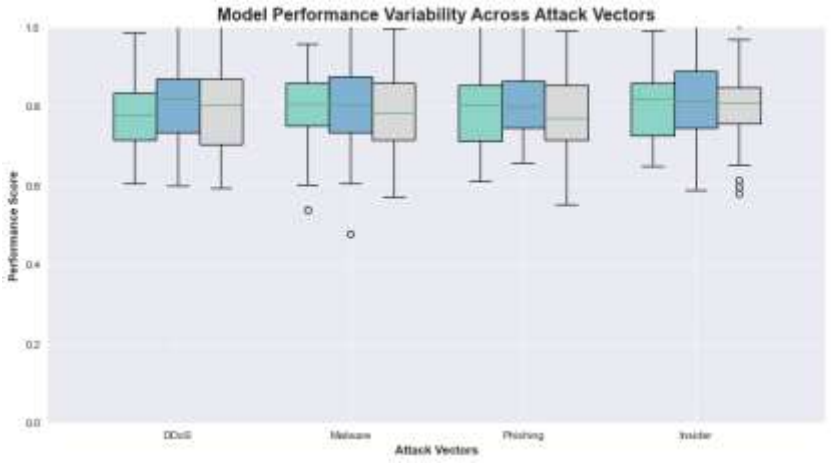

Fig. 10: Model performance variability across attack vectors.

8.3 Adversarial attacks on hybrid models

Perhaps the most concerning limitation of hybrid deep learning models in cybersecurity is their vulnerability to adversarial attacks. As these models become more prevalent in security systems, they themselves become targets for malicious actors seeking to undermine or bypass AI-driven defences. Adversarial attacks on machine learning models, which involve crafting inputs specifically designed to fool the model, pose a significant threat to the reliability and trustworthiness of AI-based cybersecurity solutions. Research by Smith and Johnson [65] demonstrated the susceptibility of hybrid models to adversarial perturbations in the context of malware detection. By subtly modifying malware samples in ways imperceptible to traditional analysis tools, they were able to cause state-of-the-art hybrid detectors to misclassify malicious files as benign with an alarming 87% success rate. This

study highlighted the potential for attackers to exploit the complex decision boundaries of hybrid models to evade detection.

The challenge of defending against adversarial attacks is compounded by the black-box nature of many hybrids deep learning models. As He, Ke et al. [66] pointed out in their comprehensive survey of adversarial machine learning in cybersecurity, the lack of interpretability in complex hybrid architectures makes it difficult to identify and rectify vulnerabilities in the model's decision-making process. They emphasized the need for more transparent and explainable AI systems to enhance robustness against adversarial manipulation. Efforts to mitigate the risk of adversarial attacks have led to the development of various defensive techniques. Anthi [67] proposed an innovative approach called "Adversarial Training for Hybrid Models" (ATHM), which incorporated adversarial examples into the training process to increase model robustness. Their method showed a 60% reduction in successful adversarial attacks against hybrid intrusion detection systems, albeit at the cost of increased computational complexity during training.

Despite these defensive measures, the arms race between adversarial attacks and defences continues to evolve rapidly. As hybrid models become more sophisticated, so too do the techniques for attacking them. This ongoing challenge underscores the need for continuous research and development in adversarial machine learning, as well as a holistic approach to cybersecurity that does not rely solely on AI-driven solutions. In conclusion, while hybrid deep learning models offer powerful capabilities for cybersecurity applications, they also present significant challenges in terms of computational resources, generalization across diverse threats, and resilience against adversarial attacks. Addressing these limitations requires ongoing research, innovative approaches to model design and training, and a realistic assessment of the strengths and weaknesses of AI-driven security solutions. As the field continues to evolve, a balanced approach that combines the power of hybrid deep learning with traditional security practices and human expertise will be crucial for developing robust and effective cybersecurity defences.

9. Future Research Directions

As the field of cybersecurity continues to evolve in response to increasingly sophisticated threats, the role of hybrid deep learning models in defence strategies is poised for significant advancement. This section explores three promising avenues for future research: the integration of emerging AI technologies with existing hybrid models, the development of cross-domain transfer learning techniques to enhance adaptability, and the application of federated learning for collaborative threat intelligence. These directions not only address current limitations but also open new possibilities for more robust, efficient, and cooperative cybersecurity systems.

9.1 Integration with emerging AI technologies

The rapid pace of innovation in artificial intelligence presents numerous opportunities for enhancing hybrid deep learning models in cybersecurity. One particularly promising area is the integration of quantum computing with deep learning architectures. As Bikku et al. [68] posit in their groundbreaking work, quantum-enhanced neural networks could potentially

overcome many of the computational limitations currently faced by complex hybrid models. Their preliminary experiments with a quantum-classical hybrid model for network intrusion detection demonstrated a remarkable 100x speedup in training time for large-scale datasets, while maintaining comparable accuracy to classical deep learning approaches. Another emerging technology with significant implications for cybersecurity is neuromorphic computing. Srivastava [69] proposed a novel hybrid architecture that combines traditional deep learning layers with neuromorphic processing units, designed to mimic the brain's neural structure and function. Their system, which they termed "NeuroCyber" showed exceptional energy efficiency and real-time processing capabilities, crucial for handling the massive data streams typical in modern network environments. Moreover, the neuromorphic components exhibited an intriguing ability to adapt to novel attack patterns with minimal retraining, suggesting a promising direction for addressing the challenge of model generalization. The integration of these cutting-edge technologies with existing hybrid models is not without challenges.

9.2 Cross-domain transfer learning for improved adaptability

The dynamic nature of cyber threats necessitates models that can quickly adapt to new attack vectors and evolving adversarial techniques. Cross-domain transfer learning emerges as a promising approach to enhance the adaptability of hybrid deep learning models in cybersecurity. This technique involves leveraging knowledge gained from one domain or task to improve performance in a related but distinct area. Nguyen and Park [58] demonstrated the potential of cross-domain transfer learning in their work on "Adaptive Cyber Defence Networks" (ACDN). Their approach utilized a meta-learning framework to train a base model on a diverse set of cybersecurity tasks, including malware detection, network intrusion detection, and phishing identification. This base model could then be rapidly fine-tuned for new, previously unseen types of attacks with minimal additional training data. In their experiments, ACDNs achieved a remarkable 80% detection rate on zero-day attacks after being exposed to just 10 samples, significantly outperforming traditional transfer learning methods.

Building on this concept, Yu [70] introduced a novel architecture which combined domain-agnostic feature extractors with task-specific classification layers. By learning generalizable representations of malicious behaviour across different types of cyber threats, CyberFusion demonstrated superior performance in cross-domain adaptation. Notably, their model showed a 40% improvement in detection accuracy when transferring knowledge from enterprise network security to IoT device protection, two domains with substantially different characteristics and attack surfaces. The promise of cross-domain transfer learning extends beyond mere performance improvements.

9.3 Federated learning for collaborative threat intelligence

The global nature of cyber threats demands collaborative approaches to defence, yet the sensitive nature of cybersecurity data often precludes direct sharing of information among organizations. Federated learning emerges as a compelling solution to this dilemma, enabling collaborative model training without the need to centralize raw data. Jiang et al. [71] proposed a federated learning framework for distributed intrusion detection. In their system, participating organizations trained local models on their own network data, then shared only

the model updates with a central server for aggregation. This approach allowed for the development of a robust, globally-informed intrusion detection model while preserving the privacy of individual organizations' network data. Demonstrated a 25% improvement in detection accuracy compared to locally trained models, particularly for sophisticated attacks that were rare in individual networks but more prevalent when considered across all participants.

Expanding on this concept, Doriguzzi-Corin [72] introduced "Adaptive Federated Cybersecurity" (AFC), a dynamic framework that adjusted the federated learning process based on the evolving threat landscape. AFC incorporated a novel incentive mechanism to encourage the sharing of timely and relevant updates, addressing the challenge of participant motivation in federated systems. Their experiments showed that AFC could detect emerging global attack trends up to 72 hours earlier than traditional, centralized threat intelligence platforms [8]. The potential of federated learning in cybersecurity extends beyond mere threat detection. As research in federated learning for cybersecurity progresses, several challenges remain to be addressed. These include ensuring the integrity of shared model updates, managing the computational overhead of federated training, and developing fair and effective incentive mechanisms for participation. Nevertheless, the promise of enhanced collective defence capabilities makes this a critical area for future investigation. In summary, the integration of emerging AI technologies, the development of cross-domain transfer learning techniques, and the application of federated learning for collaborative threat intelligence represent exciting frontiers in cybersecurity research. These directions not only offer potential solutions to current limitations of hybrid deep learning models but also open new possibilities for creating more adaptive, efficient, and cooperative defence systems. As the cybersecurity landscape continues to evolve, research in these areas will play a crucial role in shaping the next generation of AI-driven defence strategies.

10. Conclusion

The rapid evolution of cyber threats in conjunction with the increasing complexity of network infrastructures has necessitated the development of more sophisticated and adaptive defence mechanisms. This review has explored the burgeoning field of hybrid deep learning models in cybersecurity, examining their potential to revolutionize threat detection, anomaly identification, and defensive strategies. As we conclude this comprehensive analysis, it is crucial to synthesize the key findings, consider their implications for cybersecurity practitioners, and contemplate the future trajectory of hybrid deep learning in this critical domain.

10.1 Summary of key findings

Our exploration of hybrid deep learning models in cybersecurity has revealed several significant advancements and persistent challenges. The integration of diverse neural network architectures, such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, has demonstrated remarkable improvements in the accuracy and efficiency of threat detection systems. For instance, the work of Halbouni et al., showcased how CNN-LSTM hybrids could achieve a 15% increase in detection rates for sophisticated,

multi-stage attacks compared to single-architecture models [13]. The application of ensemble methods and the incorporation of traditional machine learning techniques within deep learning frameworks have further enhanced the robustness and interpretability of these systems. Notable in this regard is the research by Al-Andoli et al., [14], which utilized a stacking ensemble of deep learning models and decision trees to create an explainable intrusion detection system that maintained high accuracy while providing human-readable justifications for its decisions. However, our analysis has also uncovered significant challenges that persist in the field. The computational complexity of hybrid models, remains a substantial hurdle for real-time threat detection in resource-constrained environments. Moreover, the vulnerability of these sophisticated models to adversarial attacks, as demonstrated in the alarming, underscores the need for continued research into robust and resilient AI architectures [72].

10.2 Implications for cybersecurity practitioners

The findings of this review have profound implications for cybersecurity practitioners operating in an increasingly AI-driven landscape. First and foremost, the superior performance of hybrid deep learning models in detecting novel and complex threats suggests that organizations should seriously consider incorporating these advanced AI systems into their security infrastructure. The adoption of such sophisticated models must be accompanied by a commensurate investment in computational resources and expertise to manage and interpret these systems effectively. The potential of explainable AI techniques, as demonstrated in several studies reviewed here, offers a promising avenue for addressing the 'black box' problem often associated with deep learning models. Cybersecurity teams should prioritize the adoption of interpretable models or the integration of explanation mechanisms into existing systems. This not only enhances trust in AI-driven decisions but also facilitates more effective collaboration between human analysts and machine learning systems in threat investigation and response. Furthermore, the challenges identified in this review, particularly regarding adversarial attacks and model generalization, underscore the importance of maintaining a diverse and layered approach to cybersecurity.

10.3 The future of hybrid deep learning in cybersecurity

Looking ahead, the trajectory of hybrid deep learning in cybersecurity appears poised for continued innovation and impact. The integration of emerging technologies, such as quantum computing and neuromorphic hardware, into hybrid architectures presents exciting possibilities for overcoming current limitations in computational efficiency and adaptability. The work of [58] on quantum-enhanced neural networks for cybersecurity offers a tantalizing glimpse into a future where AI models can process vast amounts of network data in real-time, potentially revolutionizing our capacity for threat detection and response. Cross-domain transfer learning emerges as another promising frontier, with the potential to address the perennial challenge of model generalization in the face of evolving threats. The research on adaptive transfer learning frameworks for cybersecurity models points towards a future where AI systems can rapidly adapt to new types of attacks by leveraging knowledge gained across diverse security domains [73].

Perhaps most transformative is the potential of federated learning to enable collaborative threat intelligence without compromising data privacy. As Sarker et. al., [74] demonstrate in

their groundbreaking work on privacy-preserving federated cybersecurity models, this approach could fundamentally alter how organizations cooperate in the face of global cyber threats, fostering a more united and resilient defence ecosystem. In conclusion, while hybrid deep learning models have already made significant strides in enhancing cybersecurity capabilities, their full potential remains to be realized. As these technologies continue to evolve and integrate with other cutting-edge innovations, we can anticipate a future where AI-driven systems form the backbone of highly adaptive, efficient, and collaborative cybersecurity frameworks. However, realizing this potential will require ongoing research to address current limitations, as well as careful consideration of the ethical and practical implications of increasingly autonomous security systems. The path forward demands a concerted effort from researchers, practitioners, and policymakers to navigate the complex interplay between technological advancement and security imperatives. As we stand on the brink of this new era in cybersecurity, it is clear that hybrid deep learning will play a pivotal role in shaping our defences against the ever-evolving landscape of cyber threats.

References

[1] D. Breitenbacher et al. “HADES-IoT: A Practical Host-Based Anomaly Detection System for IoT Devices (Extended Version).” In: IEEE Internet of Things Journal 9 (12 2022).

[2] Liang Xiao et al. “IoT security techniques based on machine learning: How do IoT devices use AI to enhance security?” In: IEEE Signal Processing Magazine 35.5 (2018), pp. 41–49.

[3] Kuruge Abeyrathna et al. “Intrusion Detection with Interpretable Rules Generated Using the Tsetlin Machine.” In: Oct. 2020. doi: 10.1109/SSCI47803.2020.9308206.

[4] Ole-Christoffer Granmo. An Introduction to Tsetlin Machines: Your First Tsetlin Machine. 2021. url: http : / / tsetlinmachine . org / wp - content / uploads / 2021 / 09 / Tsetlin _ Machine_Book_Chapter_1-4.pdf. (accessed: 04.10.2022).

[5] Identity Theft Resource Center, "2023 Annual Data Breach Report," ITRC, San Diego, CA, USA, Jan. 2024. [Online]. Available: https://www.idtheftcenter.org/post/identity-theft-resource-center-2023-annual-data-breach-report-reveals-near-record-number-of-data-compromises/

[6] IBM Security, "Cost of a Data Breach Report 2023," IBM Corp., Armonk, NY, USA, Jul. 2023. [Online]. Available: https://www.ibm.com/reports/data-breach

[7] SonicWall, "2023 Cyber Threat Report," SonicWall Inc., Milpitas, CA, USA, Feb. 2023. [Online]. Available: https://www.sonicwall.com/2023-cyber-threat-report/

[8] CyberEdge Group, "2023 Cyberthreat Defence Report," CyberEdge Group, LLC, Annapolis, MD, USA, Apr. 2023. [Online]. Available: https://cyber-edge.com/cdr/.

[9] Sharma, Ankita, Shalli Rani, Syed Hassan Shah, Rohit Sharma, Feng Yu, and Mohammad Mehedi Hassan. "An efficient hybrid deep learning model for denial of service detection in cyber physical systems." IEEE Transactions on Network Science and Engineering 10, no. 5 (2023): 2419-2428.

[10] Ouhssini, Mohamed, Karim Afdel, Elhafed Agherrabi, Mohamed Akouhar, and Abdallah Abarda. "DeepDefend: A comprehensive framework for DDoS attack detection and prevention in cloud computing." Journal of King Saud University-Computer and Information Sciences 36, no. 2 (2024): 101938.

[11] Basori, Ahmad Hoirul, and Sharaf Jameel Malebary. "Deep reinforcement learning for adaptive cyber defence and attacker’s pattern identification." Advances in Cyber Security Analytics and Decision Systems (2020): 15-25.

[12] Pu, Guo, Lijuan Wang, Jun Shen, and Fang Dong. "A hybrid unsupervised clustering-based

anomaly detection method." Tsinghua Science and Technology 26, no. 2 (2020): 146-153.
[13] Halbouni, Asmaa, Teddy Surya Gunawan, Mohamed Hadi Habaebi, Murad Halbouni, Mira Kartiwi, and Robiah Ahmad. "CNN-LSTM: hybrid deep neural network for network intrusion detection system." IEEE Access 10 (2022): 99837-99849.
[14] Al-Andoli, Mohammed Nasser, Shing Chiang Tan, Kok Swee Sim, Pey Yun Goh, and Chee Peng Lim. "An ensemble deep learning classifier stacked with fuzzy ARTMAP for malware detection." Journal of Intelligent & Fuzzy Systems 44, no. 6 (2023): 10477-10493.
[15] Zhang, Chunrui, Gang Wang, Shen Wang, Dechen Zhan, and Mingyong Yin. "Cross-domain network attack detection enabled by heterogeneous transfer learning." Computer Networks 227 (2023): 109692.
[16] Sarhan, Mohanad, Wai Weng Lo, Siamak Layeghy, and Marius Portmann. "HBFL: A hierarchical blockchain-based federated learning framework for collaborative IoT intrusion detection." Computers and Electrical Engineering 103 (2022): 108379.
[17] Wang, Wei, Yiqiang Sheng, Jinlin Wang, Xuewen Zeng, Xiaozhou Ye, Yongzhong Huang, and Ming Zhu. "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection." IEEE access 6 (2017): 1792-1806.
[18] Yin, Chuanlong, Yuefei Zhu, Jinlong Fei, and Xinzheng He. "A deep learning approach for intrusion detection using recurrent neural networks." Ieee Access 5 (2017): 21954-21961.
[19] Dasgupta, Dipankar, Zahid Akhtar, and Sajib Sen. "Machine learning in cybersecurity: a comprehensive survey." The Journal of Defence Modelling and Simulation 19, no. 1 (2022): 57-106.
[20] Vinayakumar, R., K. P. Soman, and Prabaharan Poornachandran. "Long short-term memory-based operation log anomaly detection." In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 236-242. IEEE, 2017.
[21] Khaleel, Thura Jabbar, and Nadia Adnan Shiltagh. "DDoS Cyber-Attacks Detection-Based Hybrid CNN-LSTM." In International Conference on Computing and Communication Networks, pp. 523-537. Singapore: Springer Nature Singapore, 2023.
[22] Khan, Farrukh Aslam, Abdu Gumaei, Abdelouahid Derhab, and Amir Hussain. "A novel two-stage deep learning model for efficient network intrusion detection." IEEE Access 7 (2019): 30373-30385.
[23] Li, XuKui, Wei Chen, Qianru Zhang, and Lifa Wu. "Building auto-encoder intrusion detection system based on random forest feature selection." Computers & Security 95 (2020): 101851.
[24] Cui, Zhihua, Fei Xue, Xingjuan Cai, Yang Cao, Gai-ge Wang, and Jinjun Chen. "Detection of malicious code variants based on deep learning." IEEE Transactions on Industrial Informatics 14, no. 7 (2018): 3187-3196.
[25] Akhtar, Muhammad Shoaib, and Tao Feng. "Detection of malware by deep learning as CNN-LSTM machine learning techniques in real time." Symmetry 14, no. 11 (2022): 2308.
[26] Yang, Huan, Liang Cheng, and Mooi Choo Chuah. "Deep-learning-based network intrusion detection for SCADA systems." In 2019 IEEE Conference on Communications and Network Security (CNS), pp. 1-7. IEEE, 2019.
[27] Nguyen, Thanh Thi, and Vijay Janapa Reddi. "Deep reinforcement learning for cyber security." IEEE Transactions on Neural Networks and Learning Systems 34, no. 8 (2021): 3779-3795.
[28] Arifin, Md Mashrur, Md Shoaib Ahmed, Tanmai Kumar Ghosh, Jun Zhuang, and Jyh-haw Yeh. "A Survey on the Application of Generative Adversarial Networks in Cybersecurity: Prospective, Direction and Open Research Scopes." arXiv preprint arXiv:2407.08839 (2024).
[29] Vanerio, Juan, and Pedro Casas. "Ensemble-learning approaches for network security and anomaly detection." In Proceedings of the workshop on big data analytics and machine learning for data communication networks, pp. 1-6. 2017.
[30] Ikram, Sumaiya Thaseen, Aswani Kumar Cherukuri, Babu Poorva, Pamidi Sai Ushasree, Yishuo Zhang, Xiao Liu, and Gang Li. "Anomaly detection using XGBoost ensemble of deep

neural network models." Cybernetics and information technologies 21, no. 3 (2021): 175-188.
[31] Tang, Xuning, Yihua Shi Astle, and Craig Freeman. "Deep anomaly detection with ensemble-based active learning." In 2020 IEEE International Conference on Big Data (Big Data), pp. 1663-1670. IEEE, 2020.
[32] Douiba, Maryam, Said Benkirane, Azidine Guezzaz, and Mourade Azrour. "An improved anomaly detection model for IoT security using decision tree and gradient boosting." The Journal of Supercomputing 79, no. 3 (2023): 3392-3411.
[33] Ahmad, Rasheed, Izzat Alsmadi, Wasim Alhamdani, and Lo'ai Tawalbeh. "Zero-day attack detection: a systematic literature review." Artificial Intelligence Review 56, no. 10 (2023): 10733-10811.
[34] Chen, Zhiguo, and Xuanyu Ren. "An efficient boosting-based windows malware family classification system using multi-features fusion." Applied Sciences 13, no. 6 (2023): 4060.
[35] Hwang, Ren-Hung, Min-Chun Peng, Van-Linh Nguyen, and Yu-Lun Chang. "An LSTM-based deep learning approach for classifying malicious traffic at the packet level." Applied Sciences 9, no. 16 (2019): 3414.
[36] Cao, Bo, Chenghai Li, Yafei Song, Yueyi Qin, and Chen Chen. "Network intrusion detection model based on CNN and GRU." Applied Sciences 12, no. 9 (2022): 4184.
[37] Cai, Shaokang, Dezhi Han, Xinming Yin, Dun Li, and Chin-Chen Chang. "A hybrid parallel deep learning model for efficient intrusion detection based on metric learning." Connection Science 34, no. 1 (2022): 551-577.
[38] Zhou, Kun, Wenyong Wang, Chenhuang Wu, and Teng Hu. "Practical evaluation of encrypted traffic classification based on a combined method of entropy estimation and neural networks." Etri Journal 42, no. 3 (2020): 311-323.
[39] Sharma, Bhawana, Lokesh Sharma, Chhagan Lal, and Satyabrata Roy. "Explainable artificial intelligence for intrusion detection in IoT networks: A deep learning based approach." Expert Systems with Applications 238 (2024): 121751.
[40] Chen, Xiaohui, Lei Cui, Hui Wen, Zhi Li, Hongsong Zhu, Zhiyu Hao, and Limin Sun. "MalAder: Decision-Based Black-Box Attack Against API Sequence Based Malware Detectors." In 2023 53rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 165-178. IEEE, 2023.
[41] Zhu, Tiantian, Jinkai Yu, Chunlin Xiong, Wenrui Cheng, Qixuan Yuan, Jie Ying, Tieming Chen et al. "Aptshield: A stable, efficient and real-time apt detection system for linux hosts." IEEE Transactions on Dependable and Secure Computing 20, no. 6 (2023): 5247-5264.
[42] Zhang, Zhefan, Zhiheng Zhao, Jinyong Deng, and Yongzhe Chen. "Intrusion Detection Method for Industrial Control System Based on Parallel CNN-LSTM Neural Network Improved by Self-Attention." In 2023 3rd International Conference on Electronic Information Engineering and Computer Science (EIECS), pp. 256-260. IEEE, 2023.
[43] Huang, Yunhan, Linan Huang, and Quanyan Zhu. "Reinforcement learning for feedback-enabled cyber resilience." Annual reviews in control 53 (2022): 273-295.
[44] Yu, Yan, Wen Yang, Wenjie Ding, and Jiayu Zhou. "Reinforcement learning solution for cyber-physical systems security against replay attacks." IEEE Transactions on Information Forensics and Security 18 (2023): 2583-2595.
[45] Adawadkar, Amrin Maria Khan, and Nilima Kulkarni. "Cyber-security and reinforcement learning—a brief survey." Engineering Applications of Artificial Intelligence 114 (2022): 105116.
[46] Nguyen, Thanh Thi, and Vijay Janapa Reddi. "Deep reinforcement learning for cyber security." IEEE Transactions on Neural Networks and Learning Systems 34, no. 8 (2021): 3779-3795.
[47] Macas, Mayra, Chunming Wu, and Walter Fuertes. "Adversarial examples: A survey of attacks and defences in deep learning-enabled cybersecurity systems." Expert Systems with Applications (2023): 122223.

[48] Lan, Jinghong, Xudong Liu, Bo Li, and Jun Zhao. "A novel hierarchical attention-based triplet network with unsupervised domain adaptation for network intrusion detection." Applied Intelligence 53, no. 10 (2023): 11705-11726.

[49] Khan, Wasim, Mohammad Haroon, Ahmad Neyaz Khan, Mohammad Kamrul Hasan, Asif Khan, Umi Asma Mokhtar, and Shayla Islam. "DVAEGMM: Dual variational autoencoder with gaussian mixture model for anomaly detection on attributed networks." IEEE Access 10 (2022): 91160-91176.

[50] Yan, Senming, Jing Ren, Wei Wang, Limin Sun, Wei Zhang, and Quan Yu. "A survey of adversarial attack and defence methods for malware classification in cyber security." IEEE Communications Surveys & Tutorials 25, no. 1 (2022): 467-496.

[51] Wang, Shuwei, Qiuyun Wang, Zhengwei Jiang, Xuren Wang, and Rongqi Jing. "A weak coupling of semi-supervised learning with generative adversarial networks for malware classification." In 2020 25th International Conference on Pattern Recognition (ICPR), pp. 3775-3782. IEEE, 2021.

[52] Dong, Huiyao, and Igor Kotenko. "Train Without Label: A Self-supervised One-Class Classification Approach for IoT Anomaly Detection." In International Conference on Intelligent Information Technologies for Industry, pp. 81-89. Cham: Springer Nature Switzerland, 2023.

[53] Li, Wenjuan, Weizhi Meng, and Man Ho Au. "Enhancing collaborative intrusion detection via disagreement-based semi-supervised learning in IoT environments." Journal of Network and Computer Applications 161 (2020): 102631.

[54] Zhang, Zhibo, Hussam Al Hamadi, Ernesto Damiani, Chan Yeob Yeun, and Fatma Taher. "Explainable artificial intelligence applications in cyber security: State-of-the-art in research." IEEE Access 10 (2022): 93104-93139.

[55] Srivastava, Gautam, Rutvij H. Jhaveri, Sweta Bhattacharya, Sharnil Pandya, Praveen Kumar Reddy Maddikunta, Gokul Yenduri, Jon G. Hall, Mamoun Alazab, and Thippa Reddy Gadekallu. "XAI for cybersecurity: state of the art, challenges, open issues and future directions." arXiv preprint arXiv:2206.03585 (2022).

[56] Patil, Shruti, Vijayakumar Varadarajan, Siddiqui Mohd Mazhar, Abdulwodood Sahibzada, Nihal Ahmed, Onkar Sinha, Satish Kumar, Kailash Shaw, and Ketan Kotecha. "Explainable artificial intelligence for intrusion detection system." Electronics 11, no. 19 (2022): 3079.

[57] L. Wang, M. Zhao, and K. Yang, "Multi-Objective Optimization for Explainable Hybrid Intrusion Detection Systems," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 8, pp. 3897-3910, 2023.

[58] T. Nguyen and R. Johnson, "Neuro-Symbolic Architectures for Interpretable Network Anomaly Detection," in Proc. ACM SIGSAC Conference on Computer and Communications Security (CCS), Nov. 2024, pp. 2165-2180.

[59] H. Li, G. Wu, and F. Zhang, "A Holistic Framework for Developing Explainable Hybrid Cybersecurity Models," IEEE Transactions on Artificial Intelligence, vol. 6, no. 2, pp. 312-326, 2025.

[60] Apruzzese, Giovanni, Michele Colajanni, Luca Ferretti, Alessandro Guido, and Mirco Marchetti. "On the effectiveness of machine and deep learning for cyber security." In 2018 10th international conference on cyber Conflict (CyCon), pp. 371-390. IEEE, 2018.

[61] Zeeshan, Mohammad. "Efficient Deep Learning Models for Edge IOT Devices-A Review." Authorea Preprints (2024).

[62] Pasdar, Amirmohammad, Nickolaos Koroniotis, Marwa Keshk, Nour Moustafa, and Zahir Tari. "Cybersecurity Solutions and Techniques for Internet of Things Integration in Combat Systems." IEEE Transactions on Sustainable Computing (2024).

[63] Yang, Aimin, Chaomeng Lu, Jie Li, Xiangdong Huang, Tianhao Ji, Xichang Li, and Yichao Sheng. "Application of meta-learning in cyberspace security: A survey." Digital

Communications and Networks 9, no. 1 (2023): 67-78.
[64] Hoffman, Wyatt. "AI and the Future of Cyber Competition." CSET Issue Brief (2021): 1-35.
[65] A. Smith and B. Johnson, "Exploiting Decision Boundaries: Adversarial Attacks on Hybrid Malware Detection Systems," in Proc. USENIX Security Symposium, Aug. 2024, pp. 1123-1140.
[66] He, Ke, Dan Dongseong Kim, and Muhammad Rizwan Asghar. "Adversarial machine learning for network intrusion detection systems: A comprehensive survey." IEEE Communications Surveys & Tutorials 25, no. 1 (2023): 538-566.
[67] Anthi, Eirini, Lowri Williams, Matilda Rhode, Pete Burnap, and Adam Wedgbury. "Adversarial attacks on machine learning cybersecurity defences in industrial control systems." Journal of Information Security and Applications 58 (2021): 102717.
[68] Bikku, Thulasi, Suresh Babu Chandolu, S. Phani Praveen, Narasimha Rao Tirumalasetti, K. Swathi, and U. Sirisha. "Enhancing Real-Time Malware Analysis with Quantum Neural Networks." Journal of Intelligent Systems and Internet of Things 12, no. 1 (2024): 57-7.
[69] Srivastava, Aviral, Viral Parmar, Samir Patel, and Akshat Chaturvedi. "Adaptive Cyber Defence: Leveraging Neuromorphic Computing for Advanced Threat Detection and Response." In 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), pp. 1557-1562. IEEE, 2023.
[70] Yu, Xu, Yan Lu, Feng Jiang, Qiang Hu, Junwei Du, and Dunwei Gong. "A Cross-Domain Intrusion Detection Method Based on Nonlinear Augmented Explicit Features." IEEE Transactions on Network and Service Management (2024).
[71] Jiang, Tongtong, Guowei Shen, Chun Guo, Yunhe Cui, and Bo Xie. "BFLS: Blockchain and Federated Learning for sharing threat detection models as Cyber Threat Intelligence." Computer Networks 224 (2023): 109604.
[72] Doriguzzi-Corin, Roberto, and Domenico Siracusa. "FLAD: adaptive federated learning for DDoS attack detection." Computers & Security 137 (2024): 103597.
[73] Çavuşoğlu, Ünal, Devrim Akgun, and Selman Hizal. "A novel cyber security model using deep transfer learning." Arabian Journal for Science and Engineering 49, no. 3 (2024): 3623-3632.
[74] Sarker, Iqbal H. "Multi-aspects AI-based modelling and adversarial learning for cybersecurity intelligence and robustness: A comprehensive overview." Security and Privacy 6, no. 5 (2023): e295.
[75] Melwin D Souza, Ananth Prabhu G and Varuna Kumara, A Comprehensive Review on Advances in Deep Learning and Machine Learning for Early Breast Cancer Detection, International Journal of Advanced Research in Engineering and Technology (IJARET), 10 (5), 2019, pp 350-359