Advancing Intrusion Detection Systems: A Comprehensive Survey on AI, ML, and DL Approaches

Vigneshwar Sundararajan and Nabhan Aziz
Department of Computer Science: Cybersecurity and Privacy
University of Central Florida
Email: {vi126747, na734868}@ucf.edu

***Abstract*— The rapid evolution of cyber threats has necessitated the development of sophisticated Intrusion Detection Systems (IDS) to safeguard critical infrastructures and digital assets. Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have emerged as transformative technologies, driving significant advancements in IDS capabilities. This survey paper provides a comprehensive review of state-of-the-art AI/ML/DL methodologies applied to intrusion detection, examining their effectiveness, limitations, and potential for addressing emerging security challenges.**

By analyzing 20 seminal research contributions, this study categorizes IDS approaches based on their underlying models, evaluation metrics, and datasets, offering insights into the comparative performance of techniques such as supervised learning, anomaly detection, and hybrid systems. Additionally, the paper highlights trends, including the integration of Explainable AI (XAI), the use of adversarial robustness mechanisms, and the adoption of real-time detection frameworks. Through critical analysis, this survey identifies research gaps, such as scalability, dataset diversity, and deployment challenges, while proposing future directions to enhance the efficacy and reliability of IDS. This work aims to serve as a valuable resource for researchers and practitioners in the pursuit of resilient cybersecurity solutions.

Index Terms **Intrusion Detection Systems, AI Security, Network Security, Adversarial Machine Learning, Real-time Detection, Hybrid Models.**

I. Introduction

The ever-increasing connectivity of digital systems and the ubiquity of Internet of Things (IoT) devices have profoundly transformed the cybersecurity landscape. While these advancements have unlocked unprecedented opportunities for innovation, they have also exposed networks to a rapidly growing array of sophisticated cyber threats. Intrusion Detection Systems (IDS) have emerged as essential components in the arsenal of modern cybersecurity, designed to detect and mitigate malicious activities targeting network and host systems [1]. However, traditional IDS, which primarily rely on signature-based and rule-based techniques, often fail to address evolving threats such as zero-day attacks and adversarial intrusions [2]. This limitation underscores the need for innovative approaches to safeguard critical infrastructures.

Recent advancements in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have revolutionized the capabilities of IDS [3]. By leveraging the ability of ML algorithms to extract features from data and the hierarchical representation capabilities of DL models, AI-driven IDS can adapt to evolving attack patterns and effectively detect anomalies in large-scale networks [4]. For instance, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have demonstrated exceptional performance in classifying cyber threats [5], while hybrid models combining anomaly-based and signature-based detection offer enhanced robustness [6].

This survey paper provides a comprehensive review of state-of-the-art research in IDS, focusing on the role of AI/ML/DL methodologies in enhancing intrusion detection. Specifically, it analyzes 20 key research contributions, categorizing them by their underlying techniques, datasets, and performance metrics. The evaluation metrics employed in these studies, including accuracy, precision, recall, F1 score, detection rate, false positive rate, and computational efficiency, are critically examined to highlight the strengths and limitations of each approach [7]. These metrics not only serve as benchmarks for comparison but also guide researchers in addressing the challenges of building scalable, robust IDS frameworks.

Datasets play a crucial role in training and validating IDS models. Popular benchmarks like NSL-KDD, UNSW-NB15, and CICIDS2017 are extensively used to evaluate the efficacy of ML and DL techniques [8]. However, the generalizability of these datasets to real-world scenarios remains a persistent challenge [9]. In addition, adversarial machine learning poses significant threats to IDS, as attackers can exploit vulnerabilities in models to bypass detection [10]. This survey addresses these challenges by emphasizing the importance of robust metrics and data diversity in IDS research.

The organization of this paper is as follows: Section II provides foundational concepts of IDS and the integration of AI/ML/DL techniques. Section III outlines the methodology for selecting and categorizing the reviewed papers. Section IV presents a detailed analysis of the surveyed studies, emphasizing the evaluation metrics and datasets used. Section V discusses emerging trends, unresolved challenges, and future directions in IDS research. Finally, Section VI concludes the paper by summarizing key findings and their implications for advancing IDS technologies.

By systematically analyzing the metrics, datasets, and methodologies employed in recent IDS research, this survey aims to bridge the gap between academic innovation and practical implementation. It serves as a comprehensive resource for researchers and practitioners, paving the way for the development of more resilient and adaptive intrusion detection systems.

II. Intrusion Detection Systems and AI/ML/DL Integration

A. Intrusion Detection Systems: Concepts, Evolution, and Types

Intrusion Detection Systems (IDS) are a cornerstone of modern cybersecurity frameworks, designed to monitor and identify unauthorized access or malicious activities in networks and hosts [1]. These systems ensure the security, integrity, and availability of critical infrastructures and digital assets by flagging activities that deviate from established norms or known attack patterns.

1) Host-Based IDS (HIDS): These systems focus on analyzing activities at the host level, including file integrity checks, system call traces, and user behavior monitoring. HIDS are particularly effective in detecting localized threats but are resource-intensive and lack scalability for extensive networks [2].

2) Network-Based IDS (NIDS): NIDS monitor traffic across the network, analyzing packet headers and payloads to detect malicious activities. These systems are adept at identifying patterns indicative of distributed attacks but often face challenges in real-time data processing due to the volume of network traffic [3].

3) Signature-Based IDS: These systems rely on predefined rules or attack signatures to identify threats. While efficient against known threats, they fail to address zero-day attacks and dynamic adversarial strategies [9].

4) Anomaly-Based IDS: By employing statistical models and machine learning techniques, anomaly-based IDS detect deviations from baseline behavior. They excel in identifying unknown threats but are prone to high false positive rates [7].

5) Hybrid IDS: Integrating signature-based and anomaly-based techniques, hybrid systems combine the strengths of both approaches to improve detection accuracy and reduce false alarms [6].

B. AI/ML/DL: Transforming Intrusion Detection

The integration of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) into IDS has revolutionized the detection and prevention of cyber threats. These technologies enable IDS to process vast amounts of data, identify complex patterns, and adapt to emerging attack vectors, offering significant advantages over traditional systems [8].

1) Machine Learning (ML) in IDS: ML models, such as Support Vector Machines (SVM), Decision Trees, and Random Forests, are widely used for traffic classification and anomaly detection. These models rely on feature extraction, where domain-specific knowledge is essential to derive meaningful insights from raw data [11]. Despite their efficacy, the dependency on handcrafted features often limits their adaptability [3].

2) Deep Learning (DL) in IDS: DL models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), overcome the limitations of feature engineering by learning hierarchical data representations. CNNs, for example, excel in analyzing spatial patterns in network traffic, while RNNs are effective in capturing temporal dependencies [2]. Recent innovations, such as transforming network packets into images for analysis, have demonstrated the potential of DL to enhance detection accuracy and robustness [12].

3) Hybrid AI Models: Hybrid approaches that combine ML and DL methodologies leverage the strengths of both to improve detection performance. For instance, using Random Forests for feature selection followed by classification using DL models has shown superior results in identifying complex attack patterns [6].

C. Advantages of AI/ML/DL Integration

The integration of AI/ML/DL techniques into IDS provides several transformative benefits:

1) Enhanced Detection Accuracy: These technologies significantly outperform traditional methods in identifying both known and novel attack patterns [7].

2) Scalability and Real-Time Capabilities: AI-driven IDS can process large-scale data in real time, making them suitable for modern high-speed networks [8].

3) Adaptability to Evolving Threats: Continuous learning enables AI-based IDS to remain effective against zero-day vulnerabilities and adversarial attacks [2].

D. Challenges in Integrating AI/ML/DL

The integration of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) into Intrusion Detection Systems (IDS) has shown great promise, but several challenges hinder their widespread adoption and effectiveness. Addressing these challenges is crucial to advancing IDS technologies.

1) Dataset Limitations: The efficacy of AI/ML/DL models heavily relies on the quality, diversity, and size of the datasets used for training and evaluation. Existing benchmark datasets such as NSL-KDD, UNSW-NB15, and CICIDS2017, while widely used in the research community, often lack the complexity and variability of real-world network environments [9]. Many datasets are outdated or fail to incorporate the latest attack vectors, leaving models ill-equipped to handle emerging threats. Additionally, the class imbalance problem—where malicious activities represent only a small portion of the dataset—leads to biased models that prioritize normal traffic over detecting rare attacks [7]. The lack of publicly available datasets also limits reproducibility and comparability in IDS research, as researchers often rely on proprietary or simulated datasets that may not generalize well to real-world scenarios [8].

2) Adversarial Attacks: AI and ML models in IDS are increasingly vulnerable to adversarial manipulations, where attackers craft inputs designed to deceive the model. For example, adversarial samples can slightly perturb benign traffic to appear malicious or vice versa, bypassing detection systems [10]. This vulnerability is particularly concerning for deep learning

models, which, despite their accuracy, often act as "black boxes" with limited interpretability. Without robust defenses, adversarial attacks can compromise the reliability of IDS and expose critical systems to undetected threats.

3) Computational Overheads: Training deep learning models is resource-intensive, requiring significant computational power, memory, and time. For instance, models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks involve millions of parameters that demand advanced hardware such as GPUs or TPUs for effective training [8]. This requirement presents a barrier for organizations with limited budgets or resource-constrained environments, such as IoT networks and edge computing platforms. Optimizing computational efficiency while maintaining performance remains a key research challenge.

The application of deep learning (DL) techniques in Intrusion Detection Systems (IDS) has significantly advanced the field of cybersecurity, providing enhanced detection capabilities against complex and evolving threats. As highlighted by Macas and Wu [13], DL methods such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and autoencoders have proven highly effective in identifying both known and unknown attack patterns due to their ability to learn hierarchical data representations from raw network traffic. These approaches eliminate the need for extensive feature engineering, a limitation commonly associated with traditional machine learning models. However, the study also points out critical challenges, including the high computational demands of training and deploying DL models, as well as their susceptibility to adversarial attacks. This duality underscores the need for optimized architectures and robust defenses to fully harness the potential of deep learning in IDS. The comprehensive review provided in the study serves as a foundational reference for understanding the strengths and limitations of DL methods in cybersecurity applications.

4) Real-Time Constraints: Deploying AI-based IDS in real-time environments poses challenges due to latency and throughput requirements. High-speed networks generate vast amounts of data, making it difficult for AI models to process traffic in real-time without sacrificing accuracy or increasing false positives [12]. Models must strike a balance between computational efficiency and detection capability to be viable for real-world deployment.

E. Metrics-Driven Evaluation in IDS Research

Robust evaluation metrics are essential to assessing the performance and reliability of IDS. Metrics serve as standardized benchmarks for comparing different models, guiding researchers and practitioners in identifying the most effective approaches for specific use cases.

1) Commonly Used Metrics: Accuracy measures the proportion of correctly classified instances out of the total instances. While widely reported, accuracy alone can be misleading in cases of class imbalance, where models may achieve high accuracy by predominantly classifying normal traffic [4]. *Precision* quantifies the proportion of correctly identified positive cases (e.g., malicious traffic), while *recall* assesses the model's ability to capture all actual positive cases. Together, they offer a nuanced view of a model's detection capability [7]. The harmonic mean of precision and recall, the *F1 score* provides a balanced measure, particularly useful when evaluating models in imbalanced datasets [8]. Represents the proportion of benign traffic misclassified as malicious. *Minimizing False Positive Rate (FPR)* is crucial to reducing operational overheads and maintaining trust in the system's recommendations [2]. Also known as sensitivity, *Detection Rate (DR)* measures the proportion of correctly detected malicious traffic and is critical for evaluating an IDS's effectiveness [3].

2) Significance of Metrics in IDS Research: Metrics enable researchers to objectively compare methodologies across different datasets and experimental setups. For example, a model achieving high accuracy but with a high FPR may be unsuitable for real-world deployment [7]. They also highlight trade-offs between performance indicators. For instance, reducing FPR often comes at the cost of lowering recall, necessitating a balanced approach to optimize IDS performance [8].

3) Emerging Evaluation Trends: With the rise of Explainable AI (XAI), researchers are beginning to evaluate interpretability alongside traditional metrics, ensuring that IDS models not only perform well but also provide insights into their decision-making processes [4]. The need for domain adaptability has led to cross-dataset evaluations, where models are trained on one dataset and tested on another to assess their generalizability [9]. Real-time performance metrics, such as latency and throughput, are gaining prominence as AI-driven IDS transition from research to deployment [12].

By focusing on robust and diverse metrics, IDS research can address the practical challenges of scalability, adaptability, and operational efficiency, paving the way for next-generation intrusion detection systems.

III. Methodology

The methodology for this survey was meticulously designed to ensure a comprehensive and systematic review of Intrusion Detection Systems (IDS) research integrating Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). This section outlines the criteria used for paper selection, the categorization framework applied, and the approach to data extraction and synthesis, providing a solid foundation for the insights presented in this survey.

A. Paper Selection Criteria

The selection of papers was guided by a clear and rigorous set of criteria to ensure relevance and quality. The primary focus was on studies that explored the application of AI/ML/DL in IDS, excluding traditional approaches that lacked these technological integrations. To maintain academic rigor, priority was given to papers published in reputable journals and conferences, including IEEE Access, Elsevier, and MDPI. The timeline for inclusion was also an important factor, with an emphasis on research published between 2019 and 2024 to capture the latest

advancements in the field. Foundational papers that provided historical context and significant contributions were included regardless of their publication date.

Efforts were made to incorporate diverse perspectives by including studies that examined various IDS approaches, such as anomaly detection, signature-based systems, and hybrid methods. This diversity ensures a well-rounded understanding of the field. The initial search yielded over 50 papers, which were filtered based on these criteria, resulting in a curated selection of 20 high-impact studies for detailed analysis.

B. Categorization of Reviewed Papers

To facilitate a structured and focused review, the selected papers were categorized based on their primary contributions and areas of focus. The first category included papers that proposed or evaluated IDS architectures, covering host-based, network-based, and hybrid systems. This category provided insights into the design and operational principles of different IDS frameworks.

The second category focused on studies exploring AI/ML/DL methodologies. These papers detailed the use of models such as Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), and autoencoders, offering a comprehensive view of the techniques employed in IDS research. The third category included works centered on evaluation metrics and datasets, shedding light on how IDS performance is measured and the data used for training and testing. Lastly, the fourth category encompassed papers discussing challenges and future directions, highlighting unresolved issues such as adversarial threats, scalability, and dataset limitations, as well as proposing innovative solutions for advancing the field. This categorization provided a coherent framework for organizing and analyzing the diverse insights from the reviewed studies.

C. Data Extraction and Analysis

A systematic approach was employed to extract and analyze information from the selected papers. For each study, key details such as objectives, methodologies, evaluation metrics, datasets, and findings were documented. The analysis emphasized the unique contributions of each study while identifying common trends, recurring challenges, and gaps in the field.

The objectives and scope of the papers provided an understanding of the specific problems addressed and the context of the research. The methodologies offered insights into the AI/ML/DL techniques and IDS architectures utilized, highlighting innovative approaches and best practices. Evaluation metrics and datasets were critically assessed to gauge the robustness and generalizability of the findings. Finally, the strengths and limitations of each study were noted to provide a balanced perspective on their contributions and applicability.

D. Ensuring Comprehensive Coverage

To ensure comprehensive coverage, cross-referencing was performed among the selected papers to identify any missing studies or overlooked contributions. This iterative process helped refine the selection and ensured that the survey encompassed a wide range of perspectives and insights. By comparing and contrasting studies, unique contributions were highlighted, and redundancy was minimized, enhancing the overall coherence and depth of the review.

E. Integration of Insights

The insights derived from the reviewed papers were synthesized into a cohesive narrative that forms the basis of this survey. The integration of these insights allowed for an in-depth exploration of IDS concepts, the advantages and challenges of integrating AI/ML/DL, and the evaluation of key performance metrics. This synthesis provided a holistic view of the field, offering valuable guidance for researchers and practitioners in cybersecurity.

IV. Literature Review

This section delves into an in-depth analysis of 20 selected studies on Intrusion Detection Systems (IDS), emphasizing the performance metrics and datasets used. The review reflects the advancements in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) methodologies, highlighting their practical implications, challenges, and contributions to IDS research.

This Table I provides a comprehensive mapping between various types of Intrusion Detection Systems (IDS) and the techniques employed to enhance their performance. By categorizing the techniques based on their application to Host-Based IDS (HIDS), Network-Based IDS (NIDS), Signature-Based IDS, Anomaly-Based IDS, and Hybrid IDS, the table illustrates their suitability and effectiveness across diverse IDS implementations.

Support Vector Machines (SVMs) are widely regarded for their efficiency in anomaly detection, as they excel in identifying data points that deviate from established patterns. They perform at a medium level for HIDS, NIDS, and Hybrid IDS but exhibit high efficacy in Signature-Based IDS due to their reliance on labeled data and defined attack patterns [1]. Decision Trees, another commonly used technique, provide straightforward rule-based classifications. While effective at detecting known attack signatures (high performance in Signature-Based IDS), they show medium performance in other IDS types due to their limited ability to generalize across evolving threats [7].

Random Forests, an ensemble learning method, demonstrate medium to high performance across IDS types. Their robustness against overfitting and capacity for handling diverse datasets make them a preferred choice, particularly for Hybrid IDS, where they achieve high detection rates by integrating data from multiple sources [3]. Convolutional Neural Networks (CNNs), a deep learning technique, stand out for their ability to process spatial patterns in network traffic. CNNs achieve high performance in NIDS, HIDS, and Hybrid IDS, although their utility in Signature-Based IDS is limited due to the lack of spatial dependencies in attack signatures [12].

Recurrent Neural Networks (RNNs), designed for sequential data, are highly effective in identifying temporal patterns in

TABLE I
TECHNIQUES AND THEIR PERFORMANCE ACROSS IDS TYPES

Technique/Model	Host-Based IDS (HIDS)	Network-Based IDS (NIDS)	Signature-Based IDS	Anomaly-Based IDS	Hybrid IDS	Performance Level
Support Vector Machines (SVMs)	Medium	Medium	High	Medium	Medium	Good for anomaly detection
Decision Trees	Medium	Medium	High	Medium	Medium	Effective for basic intrusion detection
Random Forests	Medium	Medium	High	Medium	High	Good for robust comparisons
Convolutional Neural Networks (CNNs)	High	High	Medium	High	High	Best for spatial pattern recognition
Recurrent Neural Networks (RNNs)	High	High	Medium	High	High	Best for temporal patterns
Autoencoders	Medium	High	Low	Medium	Medium	Useful for feature reduction and unsupervised learning
Hybrid AI Models	High	High	High	High	High	Combines strengths of multiple approaches

network traffic, such as those found in Distributed Denial of Service (DDoS) attacks. They perform at a high level across most IDS types, particularly in Anomaly-Based and Hybrid IDS, where temporal dependencies are critical for accuracy [2]. Autoencoders, which are unsupervised learning models, are moderately effective in HIDS and NIDS due to their ability to reduce feature dimensionality and identify anomalies. However, they exhibit lower performance in Signature-Based IDS because their unsupervised nature does not leverage predefined attack patterns [14].

Hybrid AI models, which integrate multiple techniques such as Random Forests and CNNs, achieve consistently high performance across all IDS types. Their ability to combine the strengths of different methodologies allows them to address both known and unknown threats effectively, making them a versatile solution for modern IDS implementations [6]. These models leverage advanced feature selection and classification mechanisms, resulting in significant improvements in detection accuracy and robustness against adversarial attacks [10].

A. Performance Metrics in IDS Studies

Performance metrics are crucial for evaluating the feasibility and practicality of IDS in real-world scenarios, where high-speed networks and dynamic environments require systems to perform consistently under stringent conditions.

The performance metrics table II provides a comprehensive comparison of key factors influencing the efficacy of Intrusion Detection Systems (IDS) across 20 surveyed papers. Metrics such as throughput, latency, scalability, and adversarial robustness are highlighted to evaluate the practical application of each study. For instance, papers like [2] and [12] demonstrate high throughput capabilities, critical for handling large-scale networks, while [7] and [14] focus on low-latency detection, essential for real-time threat response. Scalability is consistently rated medium to high, reflecting efforts to adapt IDS to growing network demands. Adversarial robustness varies, with studies like [10] addressing challenges posed by advanced evasion techniques. This table effectively summarizes the strengths and limitations of each paper, providing insights for future IDS development.

TABLE II
PERFORMANCE METRICS ACROSS 20 SURVEYED PAPERS

Paper	Throughput	Latency	Scalability	Adversarial Robustness
[1]	-	-	Medium	Medium
[3]	-	-	Medium	Low
[2]	900 Mbps	300 ms	High	Medium
[8]	700 Mbps	-	High	High
[4]	600 Mbps	-	High	Low
[5]	-	400 ms	Medium	Medium
[9]	-	-	Medium	High
[11]	-	-	Medium	Low
[6]	500 Mbps	-	Medium	Low
[7]	-	200 ms	High	Medium
[15]	-	-	Medium	Low
[13]	-	-	Medium	Medium
[12]	1 Gbps	-	High	Medium
[14]	-	300 ms	Medium	High
[16]	-	250 ms	Medium	Medium
[17]	-	-	Medium	Medium
[18]	-	-	Medium	High
[19]	-	-	Medium	High
[20]	-	-	Medium	Medium
[10]	-	-	Medium	High

Throughput is a key metric that determines the volume of data processed by an IDS per unit time. For example, in [12], the authors proposed a packet-based sequential detection system capable of achieving gigabit-level processing rates without compromising detection accuracy. Similarly, [6], achieved enhanced throughput while maintaining low error rates.

Latency, which measures the time taken by the system to detect and respond to threats, is another critical metric. Real-time detection systems, such as the one proposed by [7], demonstrated sub-second latency while maintaining high detection accuracy. Moreover, [14] introduced a two-phase IDS that

balanced fast detection with thorough analysis, enabling low-latency processing in enterprise-scale networks.

Scalability emerged as a major focus area, reflecting the need for IDS to handle exponential growth in network traffic. Distributed frameworks proposed by [4] showed significant promise in achieving linear scalability while maintaining high detection efficacy. Techniques such as parallel processing, utilized in [8], further demonstrated the potential to handle millions of connections without significant performance degradation.

Adversarial robustness has become increasingly important with the rise of sophisticated evasion techniques. Research by [10] evaluated adversarial attacks on IDS models, revealing vulnerabilities and proposing adversarial training as a solution. Enhanced feature selection mechanisms, as discussed in [14], were instrumental in mitigating adversarial risks.

In addition to these metrics, energy efficiency and real-time detection capabilities were emphasized in studies like [8] and [16], underscoring the practical requirements for IDS deployed in IoT and edge computing environments.

B. Traditional Metrics in IDS Evaluation

While performance metrics dominate modern research, traditional evaluation metrics continue to serve as essential benchmarks for assessing the reliability and robustness of IDS models.

Accuracy, the proportion of correctly classified instances, is widely reported in IDS studies. Papers such as [1] and [19] highlighted the importance of accuracy but acknowledged its limitations, particularly in datasets with class imbalance. To address these challenges, researchers incorporated complementary metrics such as precision and recall. For instance, in [3], precision was critical for minimizing false positives, while recall ensured comprehensive detection of threats.

The F1 score, a harmonic mean of precision and recall, proved especially useful for evaluating systems trained on imbalanced datasets. Studies such as [7] and [20] emphasized the F1 score to provide a balanced assessment of detection performance. Furthermore, metrics like the false positive rate (FPR) and detection rate (DR) were explored in [18] and [14] to optimize IDS for practical deployment, balancing the need for high sensitivity with operational efficiency.

C. Datasets in IDS Research

The selection of datasets is a foundational aspect of IDS research, as it influences the model's training, evaluation, and generalization to real-world scenarios.

NSL-KDD remains a popular choice for benchmarking IDS due to its reduced redundancy and balanced class distribution. However, its limitations in representing modern attack patterns have been noted in [1] and [6], where researchers emphasized the need for updated benchmarks.

UNSW-NB15 has gained traction for its realistic traffic representation and diverse attack scenarios. Studies such as [8] and [17] leveraged this dataset to evaluate anomaly detection models, showcasing its applicability in modern networks.

CICIDS2017 provides a realistic mix of benign and malicious traffic, making it a preferred choice for evaluating hybrid and DL-based IDS models. Papers like [2] and [4] demonstrated its effectiveness in testing scalability and real-time detection capabilities.

Custom datasets address gaps in public benchmarks, as seen in [12], where a novel dataset was developed for evaluating image-based intrusion detection. Similarly, synthetic datasets used in [10] provided controlled environments for testing adversarial robustness, offering valuable insights into model vulnerabilities.

D. Comparative Insights

The comprehensive analysis of performance metrics and datasets yields several key insights. First, there is a noticeable shift in research focus toward practical performance metrics like throughput, latency, and scalability, as highlighted in [8]. Second, adversarial robustness is gaining prominence as IDS models face increasingly sophisticated threats, with studies like [10] leading the charge in this domain. Finally, the lack of diversity in benchmark datasets continue to be a challenge, necessitating innovative solutions like custom and synthetic datasets to bridge the gap, as demonstrated in [12].

E. Summary

This literature review underscores the evolving priorities in IDS research, from traditional metrics like accuracy and precision to advanced performance-oriented measures such as throughput and adversarial robustness. By systematically analyzing these aspects across 20 studies, this survey lays a strong foundation for discussing trends, challenges, and future directions in IDS development.

V. Discussion and Future Scope

The field of Intrusion Detection Systems (IDS) has seen rapid advancements, driven by the integration of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). These technologies have enabled the development of systems capable of addressing the growing complexity of cyber threats. This section delves into the emerging trends that shape IDS research, the unresolved challenges that hinder progress, and the potential directions for future investigation.

A. Emerging Trends in IDS Research

The adoption of advanced AI/ML/DL techniques is one of the most prominent trends in IDS research. These approaches have transformed traditional systems into adaptive frameworks capable of identifying patterns in network traffic and detecting anomalies in real time. Explainable AI (XAI) has emerged as a critical component in this evolution, addressing the interpretability challenges associated with complex models. Studies like [4] highlighted the importance of XAI in enhancing trust and reliability, allowing analysts to understand the rationale behind IDS decisions. By leveraging feature importance and

TABLE III
IDS DEEP LEARNING ARCHITECTURES

Reference	Focus	Summary	Strengths	Dataset	Evaluation
17	Performance Evaluation of Deep Learning Architectures for Intrusion Detection	Investigates the performance of different deep learning architectures for intrusion detection, employing various deep learning models with varying complexity levels and evaluating their performance on four large datasets.	Provides detailed explanations of the deep learning models used, including architecture and hyperparameter settings.	UNSW-NB15, CIC-IDS-2017, 5G-NIDD, FLNET2023	Accuracy, Precision, Recall, F1 Score
12	Image-Based Sequential Packet Representation for Real-Time NIDS	Proposes an image-based sequential packet representation method for real-time network intrusion detection, using the median number of packets per attack type to determine image dimensions and splitting the image data for training, validation, and testing.	Offers a detailed description of the methodology for packet-based feature extraction, image representation, and model development. Discusses considerations for data splitting and preventing model bias.	CIC-IDS2017, CIC-IDS2018	Accuracy, Precision, Recall, F1 Score, True Positive Rate, True Negative Rate, False Negative Rate, False Positive Rate
2	Deep Learning for Intelligent Intrusion Detection System with Hybrid Framework	Explores deep learning for an intelligent intrusion detection system, emphasizing a hybrid framework of network-based and host-based intrusion detection. It involves advanced text representation methods from natural language processing (NLP) for host-level events and utilizes multiple benchmark datasets for comparative experimentation.	Examines the application of advanced text representation methods from NLP and evaluates the effectiveness of these methods on multiple datasets. Highlights the pros and cons of network-based and host-based intrusion detection systems.	ADFA-LD, ADFA-WD, KDDCup 99, NSL-KDD, Kyoto, CICIDS 2017, UNSW-NB15, WSN-DS	Accuracy, precision, recall, F1-score, and confusion matrices.
5	Deep Learning for Intrusion Detection	Discusses the use of deep learning algorithms for intrusion detection, comparing them with traditional ML algorithms, and highlighting their strengths, including the ability to automatically extract features and learn complex patterns.	Provides a table summarizing various deep-learning-based IDS studies, including their datasets, results, and the authors' areas of interest.	KDD CUP99, NSL KDD, CIC IDS 2017, CSE CIC 2018	Accuracy, Detection, Classification
6	Anomaly-Based Intrusion Detection for IoT Using Convolutional Neural Networks (CNNs)	Proposes an anomaly-based intrusion detection model for IoT networks using CNNs. The model analyzes network traffic to identify deviations from normal behavior, indicating potential threats or attacks.	Discusses the model's focus on the intersection of IoT, Network Security Mechanisms (NIDS), and deep learning techniques (CNNs). Highlights the advantages of deep learning techniques for anomaly detection, particularly their ability to handle large datasets and complex patterns.	DARPA	Accuracy, Precision, Recall, F1-score, False positive/negative rates

visual explanations, XAI frameworks make the deployment of AI-driven IDS in real-world scenarios more feasible.

Hybrid methodologies that combine signature-based and anomaly-based detection are also becoming a cornerstone of IDS research. These models exploit the strengths of both approaches to achieve higher detection accuracy and lower false positive rates. For example, [20] introduced a two-stage classifier that blended heuristic methods with DL, offering superior performance in complex network environments. The increasing adoption of hybrid models underscores the need for versatile systems capable of adapting to evolving attack strategies.

Real-time detection capabilities are another pivotal area of focus. With the growing scale and speed of network traffic, the ability to detect and mitigate threats instantaneously is essential. In [12], the use of image-based representations of packet data enabled convolutional neural networks (CNNs) to achieve near-instantaneous detection. Concurrently, [16] emphasized latency reduction strategies in distributed systems, highlighting the practical benefits of real-time IDS for modern high-speed networks.

The rise of adversarial threats has propelled research into robust IDS models. Attackers often craft inputs designed to evade detection by exploiting vulnerabilities in AI-based systems. Studies such as [10] explored adversarial training and robust feature selection to enhance resilience against such attacks. Moreover, [14] presented a two-phase defense strategy that effectively countered untargeted white-box adversarial attacks, paving the way for more secure systems.

Finally, the demand for energy-efficient IDS has grown, particularly in IoT and edge environments where computational resources are limited. Research like [8] and [17] demonstrated

the viability of lightweight architectures that minimize resource consumption while maintaining high detection accuracy. These advancements are critical for ensuring the deployability of IDS across diverse platforms and scenarios.

B. Unresolved Challenges in IDS Research

Despite the significant strides made in IDS research, several challenges remain unresolved. A primary issue lies in the quality and diversity of datasets. Benchmarks like NSL-KDD and CICIDS2017 have been widely adopted, but they lack the ability to capture modern attack patterns and real-world complexities. Studies such as [1] and [7] noted that these datasets often oversimplify the dynamic nature of real-world network traffic, limiting the generalizability of trained models. Custom datasets, like those introduced in [12] and [18], offer potential solutions by addressing specific use cases. However, concerns about reproducibility and cross-environment adaptability persist.

Balancing detection accuracy with operational efficiency is another persistent challenge. High-accuracy models often demand substantial computational resources, making them impractical for resource-constrained environments like IoT networks. For instance, [19] and[17] highlighted the need for optimized architectures that achieve high detection rates without incurring excessive energy or computational costs. This trade-off underscores the importance of designing lightweight solutions that do not compromise on effectiveness.

Scalability remains a significant hurdle as networks continue to grow in size and complexity. Distributed learning techniques, as explored in [15], show promise in handling large-scale traffic by distributing the computational load across multiple nodes. However, challenges related to synchronization, fault tolerance, and real-time decision-making in distributed architectures remain largely unaddressed.

Adversarial threats pose an ongoing challenge to AI-driven IDS. Attackers constantly innovate new techniques to bypass detection mechanisms. While studies like [10] and [14] proposed robust defenses, these methods often introduce additional computational overheads and may not fully counter emerging attack vectors. This highlights the need for adaptive systems capable of proactively identifying and mitigating adversarial risks.

C. Future Directions in IDS Research

To address these challenges and capitalize on emerging opportunities, future research in IDS must focus on several key areas.

Enhancing dataset quality and diversity is paramount. Collaboration between academia and industry, as suggested by [8] and [14], could facilitate the development of standardized datasets that reflect contemporary network conditions, encrypted traffic, and multi-modal attack vectors. Such datasets would enable more robust and generalizable models, improving real-world applicability.

The development of lightweight and energy-efficient architectures is another critical priority. Techniques like model pruning, quantization, and collaborative edge-cloud frameworks, as explored in [17] and [20], offer promising avenues for reducing resource consumption while maintaining performance. These innovations are particularly relevant for IoT and edge computing environments, where constraints on power and processing capabilities are significant.

Adversarial defenses must evolve to anticipate and counteract sophisticated attack strategies. Proactive approaches, such as those combining adversarial training with explainable AI, have shown potential in studies like [4] and [10]. Further exploration of reinforcement learning and game-theoretic techniques could yield systems capable of dynamically adapting to new adversarial tactics.

Real-time detection systems must continue to innovate to handle high-speed and large-scale networks. Advances in distributed architectures, as highlighted in [16] and [15], demonstrate the potential for scalable, real-time IDS solutions. These systems must prioritize low-latency processing while maintaining high detection accuracy, for seamless integration into modern enterprise environments.

Finally, the integration of Explainable AI (XAI) into IDS design will play a pivotal role in bridging the gap between model complexity and interpretability. Studies like [7] emphasized the importance of transparency in fostering trust among security analysts. Future research should explore ways to enhance the interpretability of DL models without compromising performance.

D. Summary

The integration of AI, ML, and DL into IDS has revolutionized the field, offering unprecedented capabilities for detecting and mitigating cyber threats. While significant progress has been made, challenges such as dataset limitations, scalability, and adversarial robustness continue to impede widespread adoption. By focusing on innovative solutions and fostering collaboration, future research can unlock the full potential of IDS, enabling more secure and resilient networks.

VI. Conclusion

The evolution of Intrusion Detection Systems (IDS) represents a critical response to the growing sophistication and scale of cyber threats. This paper has presented a comprehensive survey of recent advancements in IDS technologies, with a particular emphasis on the integration of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). By analyzing key research works and their methodologies, this study offers valuable insights into the current state, challenges, and future directions of IDS research.

One of the key findings of this survey is the transformative role of AI/ML/DL in enhancing IDS capabilities. These technologies have enabled the development of systems capable of detecting complex attack patterns, identifying anomalies in network traffic, and adapting to evolving cyber threats. As highlighted throughout the paper, techniques such as hybrid detection, Explainable AI (XAI), and adversarial training have

addressed critical gaps in traditional IDS approaches. For example, XAI frameworks have improved the interpretability of black-box models, making AI-driven IDS more transparent and operationally viable [4]. Similarly, hybrid models have leveraged the strengths of both signature-based and anomaly-based detection, achieving higher accuracy and reduced false positive rates [6], [20].

Despite these advancements, significant challenges persist. The lack of diverse and realistic datasets continues to hinder the generalizability and real-world applicability of IDS. While custom datasets introduced in studies like [12] and [18] have addressed specific use cases, the field still lacks standardized benchmarks that reflect the complexity of contemporary networks and attack vectors. Moreover, scalability, computational efficiency, and adversarial robustness remain critical areas requiring further innovation [15], [19].

This survey also underscores the importance of future research focusing on lightweight and energy-efficient architectures, particularly for IoT and edge computing environments. Techniques like model compression and edge-cloud collaboration offer promising avenues for achieving robust detection with minimal resource consumption [8],[17]. Additionally, advancements in real-time detection and distributed learning frameworks will be crucial in addressing the challenges posed by high-speed, large-scale networks[16].

The implications of these findings extend beyond technical improvements. By addressing challenges like interpretability, scalability, and adversarial threats, IDS can become more accessible to a broader range of applications, from enterprise networks to IoT ecosystems. This progress will not only enhance the resilience of modern infrastructures but also foster greater trust and adoption of AI-driven cybersecurity solutions.

In conclusion, the integration of AI/ML/DL has redefined the potential of IDS, transforming them into intelligent, adaptive, and robust systems capable of addressing the complexities of modern cyber threats. While significant challenges remain, the advancements and future directions outlined in this paper provide a roadmap for continued innovation in IDS research. By building on these foundations, researchers and practitioners can develop IDS technologies that are not only effective but also scalable, efficient, and resilient, ensuring a safer digital landscape for years to come.

REFERENCES

[1] S. Mirlekar and K. P. Kanojia, "A comprehensive study on machine learning algorithms for intrusion detection system," in *2022 10th IEEE International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET-SIP)*. Nagpur, India: IEEE, October 2022, pp. 1–9.

[2] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41 525–41 550, 2019. [Online]. Available: https://doi.org/10.1109/ACCESS.2019.2895334

[3] A. Meena, D. Nigam, D. Sharma, and A. Chauhan, "Anomaly-based intrusion detection for iot: A deep learning approach," in *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. Delhi, India: IEEE, December 2021, pp. 1349–1356.

[4] M. T. Islam, M. K. Syfullah, J. Islam, H. S. Quadir, M. G. Rashed, and D. Das, "Exploring the potential: Ml vs. dl in network security with explainable ai (xai) insights," in *2023 26th International Conference on Computer and Information Technology (ICCIT)*. Cox's Bazar, Bangladesh: IEEE, December 2023, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ICCIT60459.2023.10441363

[5] S. Varshney, Shikha, S. Singhi, and B. Sharma, "Intelligent intrusion detection system using deep learning models," in *Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI)*. Kurukshetra, Haryana, India: IEEE, May 2021, pp. 787–793.

[6] M. Rele and D. Patil, "Intrusive detection techniques utilizing machine learning, deep learning, and anomaly-based approaches," in *2023 IEEE International Conference on Cryptography, Informatics, and Cybersecurity (ICoCICs)*. Princeton, NJ, USA: IEEE, November 2023, pp. 88–93.

[7] P. L. S. Jayalaxmi, R. Saha, G. Kumar, M. Conti, and T.-H. Kim, "Machine and deep learning solutions for intrusion detection and prevention in iots: A survey," *IEEE Access*, vol. 10, pp. 121 173–121 190, 2022, received 14 October 2022, accepted 30 October 2022, published 7 November 2022.

[8] F. Gutierrez-Portela, H. B. Arteaga-Arteaga, F. Almenares Mendoza, L. Calderón-Benavides, H.-G. Acosta-Mesa, and R. Tabares-Soto, "Enhancing intrusion detection in iot communications through ml model generalization with a new dataset (idsai)," *IEEE Access*, vol. 11, pp. 70 542–70 559, 2023. [Online]. Available: https://doi.org/10.1109/ACCESS.2023.3292267

[9] H. Sadia, S. Farhan, Y. U. Haq, R. Sana, T. Mahmood, S. A. O. Bahaj, and A. R. Khan, "Intrusion detection system for wireless sensor networks: A machine learning based approach," *IEEE Access*, vol. 12, pp. 52 565–52 582, 2024.

[10] A. Alotaibi and M. A. Rassam, "Adversarial machine learning attacks against intrusion detection systems: A survey on strategies and defense," *Future Internet*, vol. 15, no. 2, p. 62, 2023, accessed via Future Internet journal, February 2023. [Online]. Available: https://www.mdpi.com/article/10.3390/fi15020062

[11] S. V. Amanoul, A. M. Abdulazeez, D. Q. Zeebaree, and F. Y. H. Ahmed, "Intrusion detection systems based on machine learning algorithms," in *2021 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*. Shah Alam, Malaysia: IEEE, June 2021, pp. 282–287.

[12] J. Ghadermazi, A. Shah, and N. D. Bastian, "Towards real-time network intrusion detection with image-based sequential packets representation," *IEEE Transactions on Big Data*, 2024.

[13] M. Macas and C. Wu, "Review: Deep learning methods for cybersecurity and intrusion detection systems," *IEEE Access*, vol. 5, pp. 21 954–21 961, 2020.

[14] K. Roshan and A. Zafar, "Boosting robustness of network intrusion detection systems: A novel two phase defense strategy against untargeted white-box optimization adversarial attack," *Expert Systems With Applications*, vol. 249, p. 123567, 2024, available online 24 February 2024.

[15] A. Dandaras, B. Igor, J. Borges, and N. Doukas, "Machine learning applications for network intrusion detection systems," in *The 13th IEEE International Conference on Dependable Systems, Services and Technologies (DESSERT'2023)*. Athens, Greece: IEEE, October 2023, pp. 1–7.

[16] Z. Chen, M. Simsek, B. Kantarci, M. Bagheri, and P. Djukic, "Machine learning-enabled hybrid intrusion detection system with host data transformation and an advanced two-stage classifier," *Computer Networks*, vol. 250, p. 110576, 2024.

[17] T. Sowmya and E. A. M. Anita, "A comprehensive review of ai based intrusion detection system," *Measurement: Sensors*, vol. 28, p. 100827, June 2023, open access under CC BY-NC-ND license. [Online]. Available: https://doi.org/10.1016/j.measen.2023.100827

[18] W. H. Aljuaid and S. S. Alshamrani, "A deep learning approach for intrusion detection systems in cloud computing environments," *Applied Sciences*, vol. 14, no. 5381, 2024, accessed through Creative Commons Attribution License. [Online]. Available: https://www.mdpi.com/2076-3417/14/13/5381

[19] M. Wang, N. Yang, D. H. Gunasinghe, and N. Weng, "On the robustness of ml-based network intrusion detection systems: An adversarial and distribution shift perspective," *Computers*, vol. 12, no. 10, p. 209, 2023.

[20] R. Mohammad, F. Saeed, A. A. Almazroi, F. S. Alsubaei, and A. A. Almazroi, "Enhancing intrusion detection systems using a deep learning and data augmentation approach," *Systems*, vol. 12, no. 3, p. 79, 2024. [Online]. Available: https://www.mdpi.com/article/10.3390/systems12030079