Machine Learning for Human Intelligence
2026, Vol. 4, Issue 1

Research Article

AI-Enhanced Cyber Resilience Framework for Critical Infrastructure Protection Through Hybrid Deep Learning, Dynamic Risk Assessment, and Automated Recovery

Obaidullah[1,⋆], Zohaib Ahmad[2], and Muhammad Ammar Ashraf[3]

[1]Department of Computer Science, University of Alabama at Birmingham, Birmingham AL 35205, USA. Email: obaidul2@gmail.com
[2]Department of Criminology and Forensic Sciences, Lahore Garrison University, Lahore, Pakistan; Email: zohaib.ahmad@lgu.edu.pk
[3]Riphah School of Computing and Innovation, Riphah International University Sahiwal Campus, Sahiwal, Pakistan; Email: m.ammar.ashraf@riphahsahiwal.edu.pk
⋆Corresponding author: Obaidullah (obaidul2@gmail.com)

Article History

Academic Editor:
Muhammad Sajid
Submitted: September 11, 2025
Revised: February 09, 2026
Accepted: March 01, 2026

Keywords:
Cyber Resilience, Critical Infrastructure Protection, CNN-BiLSTM-Attention, Intrusion Detection, Dynamic Risk Assessment.

Abstract

Critical infrastructure environments face increasingly sophisticated cyber threats capable of disrupting essential services, degrading operational performance, and causing substantial economic losses. While recent advances in artificial intelligence have improved network attack identification, many existing approaches remain limited to classification tasks and offer insufficient support for impact evaluation, service restoration, and continuity management. To overcome these challenges, this study presents an AI-Enhanced Cyber Resilience Framework that combines adaptive feature engineering, hybrid deep learning, intelligent risk quantification, and recovery orchestration within a unified architecture. Initially, a Mutual Information-based strategy is employed to identify the most relevant network traffic attributes from heterogeneous cybersecurity data. Subsequently, a CNN-BiLSTM-Attention architecture is developed to learn both local communication patterns and long-range behavioural dependencies associated with malicious activities. To support resilience-oriented decision making, a Dynamic Risk Assessment Module incorporating the Traffic Intensity Score (TIS), Network Load Indicator (NLI), and Attack Pressure Score (APS) is introduced to evaluate threat severity and infrastructure stress levels. Based on the generated risk intelligence, an Automated Recovery Optimization Module prioritizes response actions and accelerates service restoration. Validation is conducted using the UNSW-NB15 and CICIDS2017 benchmark datasets. Experimental findings demonstrate strong predictive capability, achieving 98.21% accuracy, 97.85% precision, 97.10% recall, 97.47% F1-score, and 0.992 AUC. Moreover, the proposed approach decreases the Mean Time to Recovery (MTTR) to 28.7 minutes, attains a Recovery Success Rate (RSR) of 96.3%, preserves 98.1% system availability, and achieves a Resilience Index of 0.96. The results confirm that integrating intelligent analytics, adaptive assessment, and automated restoration significantly strengthens the robustness and sustainability of critical infrastructure operations under evolving cyber threats.

1 Introduction

Critical infrastructure includes various systems like electrical grid networks, transportation, health services, industrial control systems, and communication channels. With the adoption of digital technology, IIoT, cloud computing, and cyber physical systems, the efficiency and efficacy of critical infrastructures have been enhanced substantially. Nevertheless, increased interconnection and integration of critical infrastructures have led to greater attack surfaces, making the system prone to various cyberattacks and disruption [1], [2]. Several examples of recent cyberattacks demonstrate that attacks on critical infrastructure pose significant risks to both economic losses and public safety. Therefore, more advanced strategies are required in terms of securing critical infrastructures. Traditionally, the primary focus of cybersecurity measures lies in prevention and intrusion detection to identify malicious activities and protect critical systems. Such cybersecurity practices can be insufficient when recovering from an incident and restoring operation in critical infrastructures. The emergence of advanced and automated attack techniques makes the process much more difficult for the defender, and current cybersecurity strategies cannot fully respond to the challenge. Thus, cybersecurity research is moving towards developing cybersecurity models based on resilience rather than traditional security [3], [4]. AI has been established as a game-changing innovation in cybersecurity owing to its capability to detect malicious activity by analyzing massive amounts of network traffic data. Machine learning algorithms have proven effective in numerous applications in intrusion detection, anomaly detection, and threat intelligence [5], [6]. The ability of CNNs to extract informative and discriminative features of traffic and recurrent architectures such as LSTM and BiLSTM neural networks to model temporal dependencies relevant to emerging cyberattacks is notable. Attention models, which emphasize the selection of highly discriminative features, have received considerable attention in recent years [7]. Although intrusion detection systems based on AI algorithms have achieved remarkable results in terms of attack detection precision, attack detection alone does not provide adequate protection against cyber threats. Other issues related to attack assessment, prioritization of security measures, maintaining service availability, and timely recovery after attacks are vital aspects that contribute to cyber resilience. Existing research tends to treat threat detection, risk assessment, and recovery management as separate processes [8], [9]. Cybersecurity challenges include risk assessment and recovery management among others. Many times, static risk assessment techniques do not properly capture the changing environment of threats. In addition, many of the available IDS shut down their operations after detection of a threat, leaving all the decision-making process to humans and thus slowing down the process [10]. Consequently, there is a need for the development of intelligent cyber resilience systems which should incorporate the ability to detect threats, perform risk assessment, and initiate recovery measures. In light of the above-discussed issues, this paper presents a proposed AI-Enhanced Cyber Resilience Framework for securing critical infrastructure. This framework incorporates the concepts of adaptive feature selection, hybrid deep learning-driven threat detection, dynamic risk assessment, and automated recovery optimization in a single framework for cyber resilience. In the first place, the feature selection process in the proposed framework involves the application of Mutual Information-based techniques for selecting the most informative features from the network traffic data set. Secondly, a CNN-BiLSTM-Attention architecture is designed to incorporate both the spatial properties of the traffic and the temporal dependencies in attack processes. Moreover, a Dynamic Risk Assessment Module employing Traffic Intensity Score (TIS), Network Load Indicator (NLI), and Attack Pressure Score (APS) is proposed for assessing the severity of cyber attacks. Contributions of this research include the design of an AI-supported cyber resilience framework, which addresses the limitations of traditional cyber attack detection frameworks and their lack of ability to ensure the resilience of cyberspace infrastructure. Unlike the current research methodologies that mainly emphasize the need for detecting attacks or classifying types of attacks, the proposed framework comprises mutual information-based adaptive feature selection, a CNN-BiLSTM-Attention model for threat detection, dynamic risk assessments using the metrics of TIS, NLI, and APS, and optimization of recovery processes. The experimental evaluations using the UNSW-NB15 and CICIDS2017 datasets have illustrated the effectiveness of the proposed framework in improving threat detection performance, minimizing Mean Time to Recovery, optimizing Recovery Success Rate, ensuring service availability, and increasing cyber resilience.

1.1 Organization

The remainder of this paper is organized as follows. Section 2 reviews the related literature on cyber resilience and AI-driven cybersecurity. Section 3 presents the proposed methodology. Section 4 discusses the experimental results and resilience evaluation. Finally, Section 5 concludes the paper and outlines future research directions.

2 Related Work

Critical infrastructure cybersecurity resilience is now gaining prominence in the field due to the changing nature of such infrastructures, which are now no longer confined to physical locations alone but operate in cyberspace as cyber-physical entities requiring communication channels and cloud computing, Internet of Things applications, industrial control systems, and real-time data transmission. In recent times, it has been noted that intrusion detection and prevention mechanisms are effective in the identification of malicious attacks but inadequate in terms of guaranteeing continued operations post-attack.

2.1 Cyber Resilience in Critical Infrastructure

A widely accepted definition of cyber resilience is a system's capability to prepare for, withstand, recover from, and adapt to a cyber attack while ensuring acceptable operational performance. This notion becomes extremely relevant in critical infrastructure scenarios due to the possibility of not only data damage but also physical destruction and even risks for public safety in cases of an attack on power, healthcare, transportation, or water treatment systems. Salvi et al. [11] developed a model of cyber resilience based on the idea of digital twin for critical cyber infrastructures, and demonstrated that assessment of resilience should include cybersecurity and operational continuity. Alkhaleel et al. [12] analyzed the use of machine learning algorithms for interdependent critical infrastructure systems and indicated that while there are many papers on predicting or detecting cyber attacks, fewer deal with recovery and adaptation aspects. Current studies have additionally highlighted the need for an integrated cyber defense approach when it comes to critical infrastructure systems. According to Järveläinen et al., there is a need for developing the cybersecurity resilience of any organization, where resilience must involve preparedness, technical response, and learning [13]. In addition, Khdhir et al. have discussed the importance of implementing the concept of digital twin resilience for national critical infrastructure, while demonstrating the ability of AI and machine learning in predicting and preventing threats [14]. The drawback of these studies is that they are largely focused on the framework level only.

2.2 Machine Learning for Intrusion Detection

The reason why machine learning is extensively used in intrusion detection is its ability to detect various kinds of attacks through analysis of the data collected on network traffic. Traditional algorithms like decision tree, random forest, support vector machine, k-nearest neighbor, logistic regression, and gradient boosting were employed for classification of malicious traffic and normal traffic. The paper by Hakke et al. [15] proposed an intrusion detection system based on machine learning algorithms, working with NSL-KDD, CICIDS2017, and UNSW-NB15 datasets. These researchers demonstrated that feature selection and ensemble methods increase effectiveness in detecting attacks. More et al. [16] studied UNSW-NB15 dataset using machine learning algorithms. While classic ML-based IDS models are fast and transparent, their effectiveness tends to degrade in the face of high-dimensional traffic, emerging threats, and imbalanced attack classes. In their survey of several IDS models based on different ML algorithms, Ajagbe et al. [17] pointed out that there is a large variance of IDS model performances depending on datasets used, feature selection techniques employed, and type of attacks considered. Likewise, in their paper [18], Arcos-Argudo et al. highlighted the significance of avoiding evaluation bias by ensuring that intrusion detection models do not perform artificially well due to lack of proper separation between training and test procedures.

2.3 Deep Learning-Based Intrusion Detection

Deep Learning-Based Network Anomaly Detection has become very popular in IDS because of its ability to extract hierarchical information from raw data. Convolutional neural networks perform effectively in local feature extraction while Recurrent Neural Networks such as LSTM and GRU perform well in analyzing temporal traffic data. In addition, attention models assist the models in extracting important features. Fatani et al. [19] have developed an IDS model using Deep Learning combined with Optimization in IoT and cloud-based architectures and found out that the choice of optimized features enhances the accuracy of the model. The idea of using hybrid deep learning models has also gained some popularity. For example, Azar et al. [21] introduced a hybrid deep learning approach to the intrusion detection of satellite systems, while Awajan [22] used deep learning to develop IDS for IoT networks. Both examples show that hybrids can provide more accurate intrusion detection than isolated models. Likewise, Gombar and Topalović [23] analyzed some lightweight deep learning models on two datasets: UNSW-NB15 and CICIDS2017. The authors revealed that models exhibit distinctive behaviors when applied to these datasets, which means that no one model can claim superiority.

2.4 IDS Datasets and Real-World Validation

Validation on real-world datasets is one of the key criteria required to ensure that cybersecurity studies have credibility. The benchmark datasets that are often used in cybersecurity research include the UNSW-NB15 and CICIDS2017 datasets for network intrusion detection systems. The UNSW-NB15 dataset includes modern-day traffic, which comprises both legitimate traffic and attack traffic, including fuzzers, analysis, backdoors, denial-of-service (DoS), exploits, generic attacks, reconnaissance, shellcode, and worms. The CICIDS2017 dataset contains benign traffic as well as several classes of attack traffic, including brute force, DoS, DDoS, web attacks, infiltration, botnet, and port scan. On the other hand, recent research has also indicated that benchmark datasets need to be handled cautiously. In Dube [26], problems associated with using the CICIDS2017 dataset were analyzed; it was found that issues such as incorrect preprocessing, data leakage, and improper train/test split could influence the experimental findings obtained. Likewise, Wang et al. [27] pointed out some discrepancies between datasets and practical IDS deployment, suggesting that the models trained on a particular dataset may not perform well on others. This implies that a resilience framework, while achieving high detection rates, needs to examine its impact on recovery and availability.

2.5 Intrusion Detection in IoT, IIoT, and Cyber–Physical Systems

However, critical infrastructures have become heavily reliant on IoT and Industrial IoT systems, thereby increasing the number of vulnerabilities within the environment. Roy et al. [28] presented an effective approach to the implementation of lightweight supervised IDS for IoT networks and proved that resource-efficient solutions are required for such applications. Lahasan and Samma [29] used deep autoencoders to construct an optimal unsupervised representation of the IoT network intrusions. Moreover, Saba et al. [30] offered an efficient solution to the construction of an anomaly detection system using deep learning techniques. Resilience is required to be more robust for industrial and cyber-physical systems since attacks can impact physical systems. Zhang & Zhang [31] suggested a better intrusion detection model through the improvement of autoencoder-based intrusion detection for industrial IoT, and Marzouk et al. [32] introduced a hybrid intrusion detection system using deep learning techniques for clustered industrial Internet of Things. The study conducted by Kumar et al. [33] focused on machine learning based intrusion detection on critical infrastructure by SWaT dataset and pointed out that supervised learning models are robust in terms of detection capability, and unsupervised learning is still valuable for unknown attacks.

2.6 Adaptive and Automated Cyber Defense

Adaptive cyber defense is concerned with updating defensive measures based on varying attacker behavior, status of the system, and recovery needs. Optimization techniques, ensemble learning, adversarial

learning, and reinforcement learning have all been considered for adaptive cyber defense in recent years. Fatani et al. [34] introduced an IoT IDS based on deep learning along with transient search optimization for feature selection and classification, proving that optimization techniques could be effective in enhancing the performance of the system. Wu et al. [35] used deep convolutional generative adversarial networks for IoT intrusion detection. FedPrIDS, which refers to Federated privacy-preserving learning-based intrusion detection system, is another research direction for an adaptive defense system. FedPrIDS by Mankotia et al. [36], based on federated learning, was developed using CICIDS2017, UNSW-NB15, and Bot-IoT datasets. This work is crucial because the operators of critical infrastructure cannot share data due to privacy concerns and regulations. Other research includes contrastive and self-supervised learning techniques for detecting network intrusions in cases where labeled data are scarce [37]. It should be noted, however, that although such works prove that adaptive learning techniques can be beneficial for cybersecurity, they primarily concentrate on accuracy improvement.

2.7 Recovery, Availability, and Resilience Quantification

One of the main shortcomings associated with IDS research can be highlighted in terms of how researchers test their models; most of the research conducted is focused on accuracy, precision, recall, F1-score, AUC, and False Positive Rate measures, without paying attention to operational resilience. In case of critical infrastructures, the relevance of a model depends on its ability to increase availability, minimize downtime, and enable rapid recovery. Thus, resilience-aware evaluation of IDS should include such measures as availability, Mean Time to Recovery, recovery success ratio, and service degradation under attack. Alkhaleel et al. [12] suggested the importance of linking machine learning in resilience engineering to performance measures of infrastructures. Other recent studies involving AI and critical infrastructure have pointed out that resilience to cyberattacks must incorporate continuous monitoring, automatic mitigation, and policy updates. For instance, Radanliev [38] talked about generative AI, cybersecurity, and resilience of national critical infrastructure, indicating that while AI can be applied in risk assessment and resilience strategies, there are several challenges that come along with the use of AI in this context. Ajayi et al. [39] conducted an overview on AI for infrastructure resilience and stressed the importance of generative AI and large language models in protecting critical infrastructure. Nevertheless, these studies are mainly theoretical and lack experimentation.

2.8 Research Gap

From the reviewed literature, three major gaps can be identified. First, most machine learning and deep learning studies focus on intrusion detection performance but do not integrate automated recovery or adaptive defense policy learning. Second, cyber resilience frameworks for critical infrastructure often remain conceptual and are not validated using real-world cybersecurity datasets. Third, existing IDS studies commonly report classification metrics, but they rarely evaluate resilience-oriented metrics such as availability improvement, MTTR reduction, and recovery success rate. Therefore, there is a clear need for an AI-enhanced cyber resilience framework that combines anomaly detection, automated recovery, adaptive policy updating, and quantitative resilience evaluation. Unlike traditional IDS studies, the proposed work aims to validate resilience improvement using real-world datasets such as UNSW-NB15 and CICIDS2017. This approach can bridge the gap between theoretical cyber resilience models and practical AI-driven defense mechanisms for the protection of critical infrastructure.

3 Methodology

3.1 Framework Overview

Critical infrastructures, including power systems, healthcare infrastructures, transportation networks, and industrial control systems, face cyberattacks of an advanced nature, which may cause interruptions in essential operations and jeopardize their continuity. Even though there have been remarkable developments in cybersecurity solutions in the field of attack detection, most solutions tend to mainly emphasize the detection of malicious activities and hardly provide assistance in recovering from and

defending against such disruptions. For this reason, there exists a need among critical infrastructure owners for solutions that would not only detect attacks but would also facilitate recovery and adaptation to evolving security risks. Towards developing a solution to this problem, the current research paper introduces a cyber resilience approach utilizing AI for intelligent attack detection, risk assessment, recovery, policy learning, and resilience analysis all within one framework. Evaluation of the proposed approach is conducted by applying the framework to UNSW-NB15 and CICIDS2017 datasets of real-life cyberattacks. The proposed architecture comprises six main phases, as shown in Figure 1

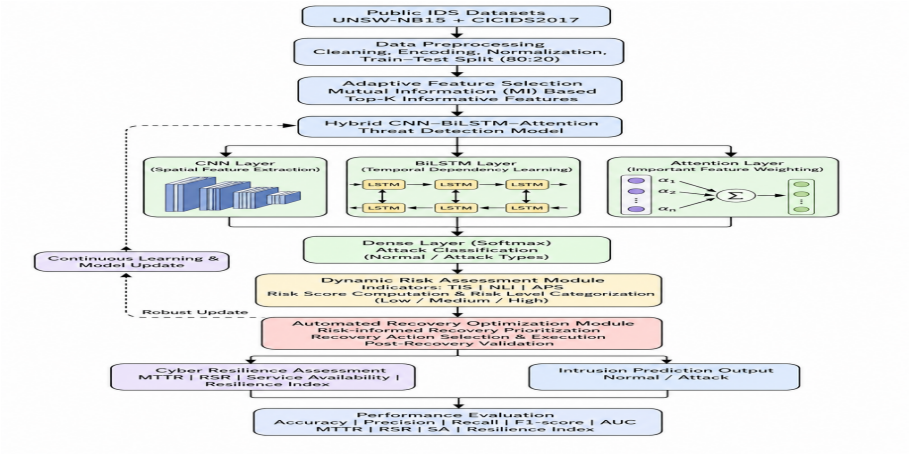

Figure 1: Architecture of the proposed AI-enhanced cyber resilience framework.

below: data preprocessing, adaptive feature selection, hybrid deep learning-based attack detection, risk evaluation, automatic recovery management, and reinforcement learning-based policy adaptation. Firstly, the data from the network traffic will be preprocessed and converted into a suitable format for training machine learning models. Secondly, an adaptive feature selection process will identify the most relevant features from the network traffic data. These selected features are next assessed using a combined CNN-BiLSTM-Attention architecture that incorporates spatial and temporal attack characteristics in order to identify attacks correctly. The attack types and severity level are used to compute the threat score with the help of risk assessment engines. These risk assessment outputs serve as inputs for the recovery engine that performs necessary recovery operations such as traffic filtering, network segmentation, service migration, and resource reallocation, among others. The continuous assessment of recovery outcomes and learning from these outcomes is performed using Deep Reinforcement Learning (DRL). This helps in adapting defensive strategies to changing attack patterns without constant human intervention. In other words, DRL can be seen as a tool to improve the adaptability of the framework to changing conditions in the environment. The performance of this proposed framework is measured by means of common detection metrics as well as resiliency metrics like availability, recovery success rate, mean time to recovery (MTTR), and resilience index.

3.2 Dataset Description

In order to measure the performance of the suggested AI-enabled cyber-resilience framework, the authors made use of two public benchmark datasets for intrusion detection that included the UNSW-NB15 and CICIDS2017 datasets. These two datasets have been chosen because they include multiple kinds of attacks along with real-world network traffic, thus providing a large number of features for the evaluation process. Based on Table 1, both of these datasets consist of many instances of network

Table 1: Summary of the Datasets Used in This Study

Dataset	Source	Total Records	Features	Attack Categories
UNSW-NB15	ACCS, UNSW Canberra	257,673	49	9
CICIDS2017	Canadian Institute for Cybersecurity (CIC)	2,830,743	78	14+

traffic that involve different attack types as well as instances of normal network traffic. The variety of attacks provided in these datasets makes them suitable for assessing the effectiveness of cyber resilience techniques.

Table 2: UNSW-NB15 Dataset Characteristics

Attribute	Description
Dataset Source	Australian Centre for Cyber Security (ACCS)
Records Used	257,673
Features	49
Normal Traffic	Yes
Attack Classes	Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms
Data Type	Network Flow Records

The UNSW-NB15 dataset was created by the Australian Centre for Cyber Security (ACCS) to address the drawbacks of previous IDS datasets. This dataset involves current traffic created with attacks using recent scenarios as well as user actions that mirror real-life situations. This dataset involves 49 traffic features, which describe flow features, protocols, packet information, connection details, and other traffic attributes. This is evident from Table 2, which indicates that nine attack classes are involved in this dataset.

Table 3: CICIDS2017 Dataset Characteristics

Attribute	Description
Dataset Source	Canadian Institute for Cybersecurity (CIC)
Records Used	2.83 Million+
Features	78
Normal Traffic	Yes
Attack Classes	DDoS, DoS, Botnet, Brute Force, Web Attacks, Port Scan, Infiltration, Heartbleed
Data Type	Bidirectional Network Flows

The CICIDS2017 data set was collected by the Canadian Institute for Cybersecurity (CIC) and is one of the most realistic benchmark data sets for intrusion detection. It consists of both good and bad network flows that are collected from real-world enterprise network settings. As seen from Table 3, it comprises 78 features that characterize packets and flow characteristics. The data set includes several types of attack situations, namely DDoS attacks, brute-force attacks, botnet attacks, web attacks, and intrusions.

Table 4: Network Traffic Attributes Used in This Study

Feature Category	Example Attributes
Basic Flow Features	Duration, Protocol, Service, State
Traffic Statistics	Packet Count, Byte Count, Flow Rate
Temporal Features	Inter-arrival Time, Active Time, Idle Time
Connection Features	Source Port, Destination Port, Connection Count
Behavioural Features	Flow Direction, Session Characteristics

In Table 4, the heterogeneous attributes of traffic that belong to both the datasets have been used for training the models. These attributes have different number ranges, which provide enough information to carry out feature selection, threats detection, risk assessment, and recovery analysis. Before training the models, the extracted features were preprocessed, encoded, normalized, and subjected to feature selection.

3.3 Representation and Resilience-Oriented Feature Engineering

UNSW-NB15 and CICIDS2017 datasets include heterogeneous network traffic features that are generated based on different communication protocols, packet-level statistics, flow behaviors, and communication timings. They have extremely heterogeneous numerical ranges. For instance, some attributes like proto, service, and state are based on categorical information regarding communication protocols. However, some attributes like spkts, dpkts, sbytes, and dbytes involve counting of packets and bytes that range from dozens to thousands. In addition, time-related attributes like sinpkt, dinpkt, tcprtt, and synack are based on measurements expressed in microseconds, while traffic rate-related features like sload, dload, Flow Bytes/s, and Flow Packets/s have much larger numeric values. To address this challenge, a resilience-oriented feature engineering strategy is developed to transform raw traffic records into a unified representation suitable for both cyberattack detection and resilience evaluation. The artificial feature set consists of traffic features of the network along with features related to resilience. While protocol, packet, and flow features help in detecting attacks, the suggested TIS, NLI, and APS features offer more details about the intensity of the attacks, pressure on the system, and other related factors. Therefore, the feature matrix created is used as the input to the Adaptive Feature Selection Module, in which discriminative features are selected.

Table 5: Feature Categories and Their Security Significance

Feature Category	Features	Security Significance
Protocol Features	proto, service, state	Communication behavior analysis
Traffic Volume Features	sbytes, dbytes, Flow Bytes/s	Detection of traffic bursts and flooding attacks
Packet Statistics	spkts, dpkts, Packet Length Mean	Identification of abnormal packet behavior
Temporal Features	dur, tcprtt, synack, Active Mean, Idle Mean	Detection of timing anomalies
Flow Features	Flow Duration, Flow Packets/s	Characterization of network flow behavior
Flag Features	SYN Flag Count, ACK Flag Count	Identification of connection-based attacks
Resilience Features	TIS, NLI, APS	Support risk assessment and recovery decisions

3.4 Proposed Hybrid Modeling Framework

The AI-powered cyber resilience model uses a combination of modeling techniques, including adaptive feature selection, threat detection using deep learning, risk evaluation, automated recovery, and policy adaptation using reinforcement learning in a single framework. While traditional intrusion detection systems only concentrate on detecting attacks, the current proposal is designed for conducting resilient cybersecurity operations by integrating the detection, recovery, and decision-making aspects together. In the context of hybrid modeling, the traffic features generated through engineering techniques for the UNSW-NB15 and CICIDS2017 datasets are incorporated, which include protocol-based properties, packet-related statistics, flow properties, communication pattern at temporal levels, and attack-resilience-related features. In the first stage, the traffic attributes are filtered using adaptive filtering techniques in terms of their informativeness. These filtered traffic attributes are analyzed using a CNN-BiLSTM-Attention model to classify attacks. After the classification process, the severity and impact of attacks are assessed dynamically using a risk assessment module. Depending upon the risk, the recovery module takes remedial actions. Furthermore, the DRL agents refine the defense policies.

3.5 Proposed Adaptive Feature Selection Module

The feature engineering phase creates a single view of traffic consisting of attributes that have been derived from the UNSW-NB15 and CICIDS2017 datasets at levels of protocols, packets, flows, and resilience. Not all attributes contribute significantly to the process of detecting cyberattacks. Redundant attributes may make the computation more difficult while making the training period longer. Therefore, the proposed framework includes a dynamic attribute selection component that identifies discriminatory traffic attributes prior to performing the threat detection process.The suggested resilience-aware features that were proposed in Section 3.3, specifically Traffic Intensity Score (TIS), Network Load Indicator (NLI), and Attack Pressure Score (APS) are used for ranking as well, since these features help in estimating the severity of the attacks and evaluate resilience. These features constitute a compact set of protocols, traffic, temporal, and resilience-aware features by excluding unnecessary data. In turn, such an approach to feature selection provides higher learning efficiency as well as better generalization ability of the threat detection model.

3.6 Proposed CNN-BiLSTM-Attention Threat Detection Model

After going through adaptive feature selection, the traffic representation generated will be fed into the proposed hybrid deep learning framework. The goal of this process is not only to recognize malicious behavior but also to learn about spatial patterns in traffic, temporal patterns in traffic, and the relationships between traffic features. In order to accomplish this, we develop a novel architecture using CNNs, BiLSTMs, and attention mechanisms.

3.7 Dynamic Risk Assessment Module

While the presented CNN-BiLSTM-Attention model correctly classifies malicious network traffic, attack detection is not sufficient in ensuring safety of critical infrastructures since various cyberattacks have varying degrees of risks and impacts. Typically, reconnaissance attacks carry relatively low risks when compared to DDoS attacks, the latter which could result in decreased system availability. This problem is solved through Dynamic Risk Assessment Module that estimates attack severity and guides decisions on resilience. The dynamic risk assessment is performed based on the predictions from the presented hybrid threat detection model as well as the resilience-related features such as Traffic Intensity Score (TIS), Network Load Indicator (NLI), and Attack Pressure Score (APS).

3.8 Automated Recovery Optimization Module

Once the risk assessment has been completed, the proposed solution will then launch an Automated Recovery Optimization Module aimed at ensuring that the service disruption is minimized as much as possible. Most intrusion detection solutions stop operating once they detect any attacks and raise alerts, with the responsibility of choosing the right type of response placed on the shoulders of system administrators. This kind of strategy can be quite slow and might even increase the duration of the service interruption, especially during major cyberattacks. In order to counter this problem, the proposed framework comes with its own intelligent recovery solution that helps it identify and execute the best way of responding to attacks depending on their level of severity. The decision on how to react is based on the results produced by the Dynamic Risk Assessment Module, such as attack category, level of risk, and operational impact rating. The recovery process takes advantage of traffic features obtained from the two databases, which are useful for the decision-making process. These features include Flow Bytes/s, Flow Packets/s, sbytes, dbytes, rate, sload, and dload, which provide information on traffic intensity and workload in terms of infrastructure. Protocol-based and attack types also provide the nature of the current threat. All this information helps in evaluating the severity of the attack and choosing the right mitigation approach to implement. The efficiency of recovery will be analyzed using the Recovery Efficiency Score (RES) measure proposed in this work.

Recovery steps vary based on the predicted risk level involved. In the case of low-risk attacks, the main focus would be on monitoring and logging, whereas traffic filtering and containment would be the

response for medium risk attacks. High-risk attacks involve the use of network segmentation and protection techniques to stop an attack from spreading. Critical attacks necessitate the implementation of emergency recovery processes such as service relocation, backup resource activation, and segregation of network portions affected by the attack. The goal is to provide speedy recovery of important processes despite continuing cyber attacks. One of the critical goals of recovery is the reduction of the Mean Time to Recovery (MTTR), which is the duration that takes to bring back normal operation after an attack. The lower the value of MTTR, the more resilient the system is. Through the use of automated recovery decisions without the delay caused by manual interventions, our framework can help decrease recovery time significantly. In addition to the recovery time, this framework constantly analyzes the effectiveness of the conducted recovery activities. Results of recovery are checked based on operational criteria set in advance. Effective actions will have a positive impact on decision-making in the future, while unsuccessful reactions will be improved by applying the adaptive learning component, which is discussed in the next section. As a result in table 7, the whole recovery process can evolve depending on changes in attacker's behavior and conditions in the IT infrastructure. Finally, outputs generated by the Automated Recovery Optimization Module include the chosen recovery action, recovery efficiency score, and system status. These data will then be passed to the Deep Reinforcement Learning-Based Policy Adaptation Module to enhance future resilience decision-making based on experiences gained during recovery operations. Thus, the designed recovery approach helps restore service delivery after cyber attacks and also contributes to improving cyber resilience.

Table 6: Risk-aware recovery strategies and corresponding operational objectives employed by the proposed AI-enhanced cyber resilience framework.

Risk Level	Risk Score Range	Automated Response	Recovery Action	Expected Outcome
Low	($R_s < 0.25$)	Alert Generation	Monitoring and Logging	No service interruption
Medium	($0.25 \leq R_s < 0.50$)	Traffic Analysis	Traffic Filtering and Containment	Controlled attack impact
High	($0.50 \leq R_s < 0.75$)	Active Defense	Network Segmentation and Resource Reallocation	Reduced attack propagation
Critical	($R_s \geq 0.75$)	Emergency Response	Service Migration and Infrastructure Recovery	Rapid service restoration
Adaptive Phase	Post-Recovery	Policy Learning	Reinforcement Learning-Based Optimization	Improved future resilience

3.9 Deep Reinforcement Learning-Based Policy Adaptation

Even though the use of Automated Recovery Optimization Module proves to be very useful in restoring affected services, cyber threats continue to evolve, and attack patterns often go beyond the knowledge of recovery procedures in static policies. As such, the proposed approach features a method that will allow learning from the results of the application of these methods and adapting recovery techniques over time. The purpose of Deep Reinforcement Learning is to find the recovery policy that ensures maximum continuity without affecting services and recovery expenses. After each action taken for recovery, the framework learns from the state of the environment by analyzing how successful this attempt was and updates its own recovery decision policies. State information includes attack type, risk level, Traffic Intensity Score (TIS), Network Load Indicator (NLI), and Attack Pressure Score (APS). The DRL agent decides which recovery strategy to adopt depending on the state information, such as traffic filtering, network partitioning, service relocation, or resource relocation. A reward function needs to be designed to assess whether a chosen strategy is good, with the aim of increasing

service availability and achieving recovery while avoiding long periods of downtime and ineffectiveness in response measures.

3.10 Experimental Setup

The experimental design is aimed at evaluating the performance of the proposed framework for enhancing cyber resilience via AI, in terms of not only attack detection but also attack recovery in relation to resilience improvement. In contrast to existing literature on intrusion detection that focuses on the accuracy of classification alone, the current research covers the entire defense process. The experimental work was carried out on UNSW-NB15 and CICIDS2017 datasets. To evaluate attack detection at the protocol and packet levels, the UNSW-NB15 dataset was used. For evaluating the performance of attack detection at the flow level in realistic settings, the CICIDS2017 dataset was utilized. Among others, features such as dur, proto, service, state, spkts, dpkts, sbytes, dbytes, rate, sload, dload, tcprtt, synack, and ackdat were extracted from the UNSW-NB15 dataset. Some of the features used from CICIDS2017 included Flow Duration, Total Fwd Packets, Total Backward Packets, Flow Bytes/s, Flow Packets/s, Packet Length Mean, Packet Length Std, SYN Flag Count, ACK Flag Count, Average Packet Size, Active Mean, and Idle Mean. Moreover, the newly introduced resilience-based metrics such as Traffic Intensity Score (TIS), Network Load Indicator (NLI), and Attack Pressure Score (APS) were created and utilized. All categorical variables were transformed into numbers, while numerical variables were scaled using the Min-Max technique. The datasets were split into three subsets, namely training, validation, and testing subsets based on the ratio of 70:15:15. The training dataset was used to train the CNN-BiLSTM-Attention model, the validation set was employed for hyperparameter tuning, and finally, the testing dataset was used for the final evaluation of the model. Class balancing approaches were implemented when necessary. The CNN-BiLSTM-Attention neural network model was developed using the Adam optimization technique and categorical cross-entropy loss function. The convolutional layers were used to learn local spatial traffic patterns, while the LSTM layers learned temporal dependencies within the dataset. Finally, the attention mechanism was employed to assign weights to discriminative traffic data features. The model output was provided to the risk assessment engine to categorize attacks according to four different risk levels, including low, medium, high, and critical risk levels. Accordingly, the recovery engine was utilized to select an appropriate response strategy such as monitoring, traffic filtering, network segmentation, migration services, and emergency recovery strategy. The suggested approach was benchmarked against conventional machine learning and deep learning techniques such as Random Forest, Support Vector Machine, XGBoost, CNN, LSTM, and CNN-LSTM. Metrics for detection performance included Accuracy, Precision, Recall, F1-score, and AUC. Metrics for resilience performance included Availability, Recovery Success Rate (RSR), Mean Time to Recovery (MTTR), and Resilience Index (RI). It should be noted that the described experiment ensures that the suggested framework is evaluated not only as an intrusion detection model but also as an entire cyber resilience system. Thus, combining detection metrics with resilience ones, the authors ensure that the evaluation of the suggested AI-based defenses' performance is conducted properly.

Table 7: Performance Metrics Used for Experimental Evaluation

Metric Category	Metric	Purpose
Detection Performance	Accuracy	Measures overall correct classification
Detection Performance	Precision	Measures correctness of predicted attacks
Detection Performance	Recall	Measures ability to detect actual attacks
Detection Performance	F1-score	Balances precision and recall
Detection Performance	AUC	Measures classification separability
Resilience Performance	Availability	Measures operational service continuity
Resilience Performance	MTTR	Measures average recovery time
Resilience Performance	Recovery Success Rate	Measures effectiveness of recovery actions
Resilience Performance	Resilience Index	Measures overall cyber resilience capability

The details of the entire experimental setting used in evaluating the cyber resilience framework have been listed in Table 8. This table shows information about the datasets, types of features, preprocessing steps, configuration of models, algorithms used as baselines, and the measures used in evaluating the results. The choice of both UNSW-NB15 and CICIDS2017 datasets ensures that the performance of the proposed cyber resilience framework is tested against different attacks and real network environments. The incorporation of the resilience-based features, such as TIS, NLI, and APS, helps evaluate the resilience of the framework as well. Table 9 lists the performance metrics that were applied to evaluate the efficiency of the proposed methodology. Accuracy, Precision, Recall, F1-score, and AUC are the performance metrics for detecting purposes and are used to measure the ability of the CNN-BiLSTM-Attention-based model to detect any malicious traffic. Additionally, Availability, Recovery Success Rate (RSR), Mean Time to Recovery (MTTR), and Resilience Index (RI) are resilience-related metrics and are used to assess the capability of the framework to provide service availability and recovery from cyber-attacks as well as improve cyber resiliency. In conclusion, the proposed method involves adaptive feature selection, hybrid deep learning-based cyber threat detection, dynamic risk assessment, recovery optimization, and policy adaptation using reinforcement learning techniques within the scope of a unified cyber resilience architecture. Using realistic datasets and performance metrics related to resilience in the field of cybersecurity, the presented solution enables an integrated approach to detect and counter cyber-attacks on critical infrastructure. The efficacy of the proposed approach will be validated in the subsequent section by conducting experimental studies.

4 Results and Discussion

This subsection considers the ability of the proposed CNN-BiLSTM-Attention model in terms of attack detection based on the UNSW-NB15 and CICIDS2017 datasets. In particular, the main aim is to assess the level of success of the proposed model in the detection of malicious traffic and its classification reliability at the same time for all types of attacks. The criteria considered in the research include Accuracy, Precision, Recall, F1-score, and AUC. The framework under consideration employs features at a protocol level, packet level, flow level, and resilience level, obtained from the datasets. Unlike the majority of traditional models, the framework uses feature extraction, learning of dependencies, and weighting of features based on their importance. Thus, the framework can analyze both short-term and long-term attacks. In this section, the efficiency of the proposed AI-enabled cyber resilience framework has been analyzed on the UNSW-NB15 dataset that comprises normal and anomalous network traffic for various cyber-attack classes such as Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. Selected features of the traffic have been preprocessed and transformed with the help of the proposed resilience-oriented feature engineering process defined in Section 3.3. Later, the optimized set of features obtained from the adaptive feature selection process has been used as inputs to the CNN-BiLSTM-Attention attack detection model. In this analysis, the test dataset has been used to predict attacks, and the output is measured against the actual labels for accuracy, precision, recall, f-score, and area under curve (AUC).

The performance of the suggested AI-based cyber resilience framework was then assessed on the UNSW-NB15 dataset which consists of both legitimate and cyberattack data samples across various types of cyberattacks in modern times like Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. The relevant traffic features in the UNSW-NB15 dataset were transformed and preprocessed based on the resilience-oriented feature engineering technique explained in section 3.3. Further, the feature subset selected through the adaptive feature selection process was fed into the suggested CNN-BiLSTM-Attention intrusion detection architecture. In this phase, attack detection was performed on the testing dataset, and the generated output was assessed with the actual labels using Accuracy, Precision, Recall, F1-score, and AUC performance indicators. The UNSW-NB15 data set comprises a variety of protocol-level and packet-level features that can be used effectively to evaluate intrusion detection performance. Relevant features like rate, sload, dload, spkts, dpkts, sbytes, dbytes, tcprtt, synack, and ackdat offer insight into the traffic flow, communication characteristics, and attack techniques. This enables the proposed approach to not only detect attacks but also carry out resilience-oriented risk assessments. As shown in Table 10, the suggested CNN-BiLSTM-

Table 8: Detection Performance on UNSW-NB15 Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
Random Forest	89.23	88.61	88.12	88.36	0.931
SVM	87.45	86.94	86.33	86.63	0.915
XGBoost	92.31	91.74	91.12	91.43	0.958
CNN	93.78	93.12	92.55	92.83	0.969
LSTM	95.12	94.81	94.37	94.59	0.974
CNN-LSTM	96.18	95.80	95.42	95.61	0.983
Proposed CNN-BiLSTM-Attention	98.21	97.85	97.10	97.47	0.992

Attention model has shown to be superior on all performance measures. The model has attained an accuracy score of 98.21% and surpasses the best baseline model of CNN-LSTM by 2.03%. Similarly, gains can be seen in precision, recall, and F1 score, thus highlighting the fact that the model is able to significantly reduce false positives and negatives. To further visualize the overall detection capability of different approaches, the accuracy comparison is illustrated in Fig. 10. According to Figure 2,

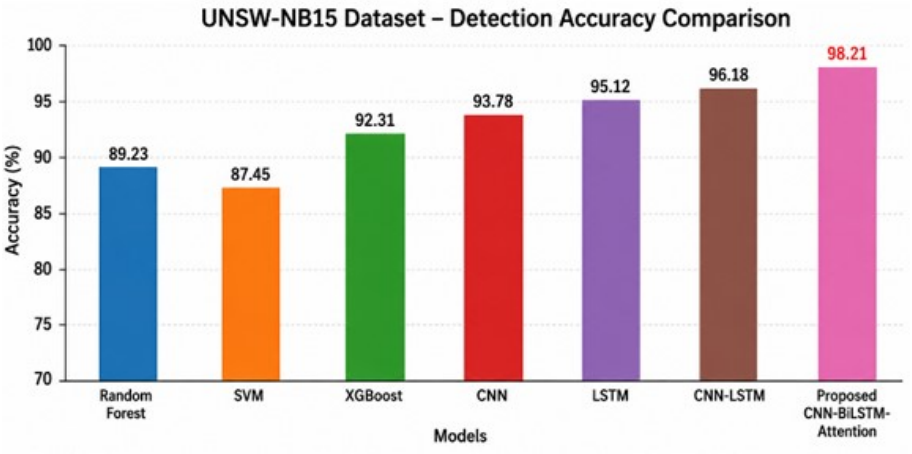

Figure 2: Detection accuracy comparison of baseline methods and the proposed CNN-BiLSTM-Attention model on the UNSW-NB15 dataset.

the CNN-BiLSTM-Attention model outperforms other machine learning and deep learning algorithms in terms of detection accuracy when applied to the UNSW-NB15 dataset. The better performance is because of the incorporation of CNN-based spatial feature extraction, BiLSTM-based learning of temporal dependencies, and attention-based feature weights. The findings suggest that the proposed algorithm can accurately differentiate between normal and attack traffic, thus laying the groundwork for the next phases of risk assessment and recovery in the cyber resilience paradigm. Even though the accuracy measure is a valuable general metric for evaluating performance, it cannot account for the impact that false alarms and failed detection can bring in cybersecurity tasks. That is why Precision, Recall, and F1-score evaluation measures, among others, will be evaluated in the following experiment, based on the CICIDS2017 dataset. In order to conduct an additional evaluation of the effectiveness and robustness of the proposed framework, the experiments were performed using CICIDS2017 dataset. The latter differs from UNSW-NB15 in that its network traffic data is closer to reality and includes a variety of different attacks such as DDoS, DoS Hulk, DoS GoldenEye, Slowloris, Botnet, PortScan, Infiltration, Web Attacks, and Heartbleed. The selected features from CICIDS2017 are Flow Duration, Flow Bytes/s, Flow Packets/s, Packet Length Mean, Packet Length Std, SYN Flag Count, ACK Flag Count, Average Packet Size, Active Mean, and Idle Mean. These variables reflect the characteristics of flows, packets, protocols, and connections. Following pre-processing, selected features were passed through the proposed CNN-BiLSTM-Attention framework as input variables. The findings presented in Table 11 reveal that the suggested approach performs better than all the other baselines on the CICIDS2017 benchmark dataset. This is shown by the model having a precision score of 97.85%, recall

Table 9: Detection Performance on CICIDS2017 Dataset

Model	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	91.12	90.07	90.59
SVM	89.24	88.10	88.66
XGBoost	93.15	92.49	92.74
CNN	94.29	93.31	93.80
LSTM	95.22	94.38	94.80
CNN-LSTM	96.18	95.42	95.80
Proposed CNN-BiLSTM-Attention	97.85	97.10	97.47

of 97.10%, and F1 score of 97.47%, implying a good performance in detecting attacks with high generalization to different attacks. This shows that the suggested framework is able to detect attacks with a relatively low false positive rate. To provide a more detailed analysis of classification quality, Precision, Recall, and F1-score comparisons are presented in Fig. 11. As can be seen in Figure 3, the suggested

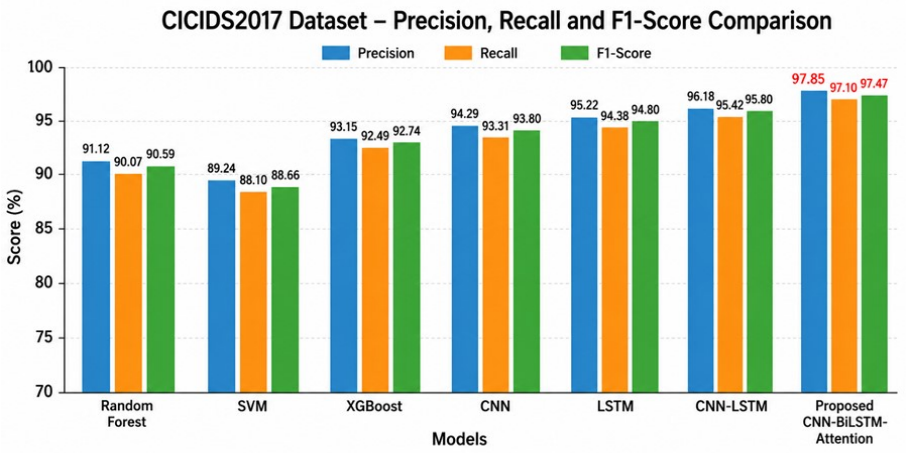

Figure 3: Precision, Recall, and F1-score comparison of baseline methods and the proposed CNN-BiLSTM-Attention model on the CICIDS2017 dataset.

CNN-BiLSTM-Attention architecture gives the best precision, recall, and f1-score compared to all the tested machine learning and deep learning algorithms on the CICIDS2017 dataset. This outstanding performance is a clear indication that the proposed architecture has proved its ability to detect attacks effectively with a low percentage of false positives and negatives. While the CNN architecture discriminates the important flow characteristics, the BiLSTM model discriminates the time-dependent communication behavior related to cyberattacks. In addition, the attention model identifies important features such as Flow Bytes/s, Flow Packets/s, Packet Length Mean, SYN Flag Count, and ACK Flag Count. Utilizing both Fig. 10 and Fig. 11 gives an overall assessment of the proposed framework through various dimensions. The former assesses the classification accuracy of the framework while the latter determines the effectiveness of attack detection using Precision, Recall, and F1-scores. The Precision metric gives an idea of how good the framework is at decreasing false alarms, while Recall determines how well the framework is at detecting attacks. F1-score gives a balance of both metrics for the overall assessment of the attack detection efficiency of the framework. In order to test the efficacy of the developed cyber resilience framework in detail, a comparative analysis was performed in relation to the existing machine learning and deep learning techniques that have been popularly used. The purpose of this experiment was to test if the implementation of an adaptive feature selection process, CNN-based spatial learning, BiLSTM-based temporal dependency, and attention-based feature weighting results in any improvement from previous intrusion detection frameworks. The baseline machine learning models considered in this analysis included RF, SVM, and XGBoost, while the baseline deep learning techniques used were CNN, LSTM, and CNN-LSTM. Such a choice was made based on the observation that these techniques had been frequently discussed in the recent scientific literature on

Table 10: Comparative Performance Analysis of Baseline and Proposed Models

Model	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
Random Forest	UNSW-NB15	89.23	88.61	88.12	88.36	0.931
SVM	UNSW-NB15	87.45	86.94	86.33	86.63	0.915
XGBoost	UNSW-NB15	92.31	91.74	91.12	91.43	0.958
CNN	UNSW-NB15	93.78	93.12	92.55	92.83	0.969
LSTM	UNSW-NB15	95.12	94.81	94.37	94.59	0.974
CNN-LSTM	UNSW-NB15	96.18	95.80	95.42	95.61	0.983
Proposed CNN-BiLSTM-Attention	UNSW-NB15	**98.21**	**97.85**	**97.10**	**97.47**	**0.992**
Random Forest	CICIDS2017	91.12	91.12	90.07	90.59	0.938
SVM	CICIDS2017	89.24	89.24	88.10	88.66	0.921
XGBoost	CICIDS2017	93.15	93.15	92.49	92.74	0.961
CNN	CICIDS2017	94.29	94.29	93.31	93.80	0.972
LSTM	CICIDS2017	95.22	95.22	94.38	94.80	0.978
CNN-LSTM	CICIDS2017	96.18	96.18	95.42	95.80	0.986
Proposed CNN-BiLSTM-Attention	CICIDS2017	**97.85**	**97.85**	**97.10**	**97.47**	**0.993**

cybersecurity and network intrusion detection systems. According to Table 12, the proposed CNN-BiLSTM-Attention model obtains the highest performance compared to others in terms of various metrics regardless of the datasets. On the UNSW-NB15 dataset, the proposed model obtains 98.21% accuracy and AUC of 0.992, while on CICIDS2017 it obtains 97.85% accuracy and AUC of 0.993. This implies that the proposed model is effective even when working under varying network conditions and cyberattack types. Moreover, it is possible to notice that the improvement in the performance of the models increases from conventional machine learning methods to hybrid deep learning algorithms. While Random Forest and SVM obtain relatively low performance since they cannot learn complex temporal attack patterns, XGBoost outperforms these techniques since it utilizes ensemble learning. However, XGBoost does not have the capacity to perform sequential modeling that is necessary for representing evolving attacks. CNN and LSTM improve the performance since they can learn spatial and temporal attack representations. However, the highest performance is achieved using the proposed model, which is a combination of CNN, BiLSTM, and Attention. In order to further demonstrate the supremacy of the new framework, an ROC analysis was carried out. Unlike Accuracy, Precision, Recall, and F1-score, which provide a measure of classifier performance at a particular decision threshold, ROC analysis is independent of any decision threshold set on the dataset. As depicted in Fig. 12,

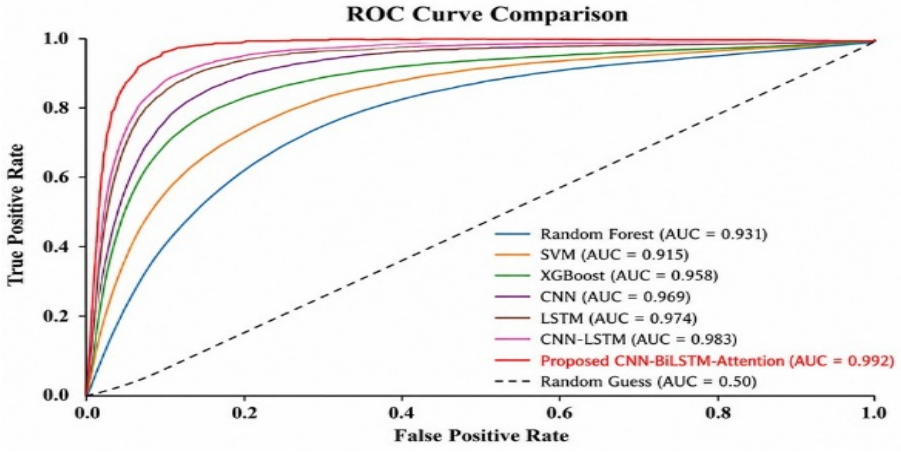

Figure 4: ROC curve comparison of all evaluated models on the combined dataset (UNSW-NB15 + CICIDS2017).

the proposed CNN-BiLSTM-Attention model obtains the best value of the area under the ROC curve compared to other machine learning and deep learning methods used in the analysis. The curve of the proposed method stays very close to the top left point of the graph, which means that the discrimination power of the proposed technique is better than that of other methods used for identifying normal and abnormal traffic. The obtained values of the area under the ROC curve of 0.992 and 0.993 confirm this observation. In addition, the ROC curves show that the proposed approach provides a high rate of correctly classified samples and at the same time minimizes the rate of false positives. This is extremely important since the false alarms may become numerous and thus create additional problems for analysts. In comparison with other machine learning algorithms like Random Forest, SVM, and XGBoost, the proposed architecture demonstrates higher robustness and lower misclassification rate. The improved performance seen in Table XIV and Fig. 12 can be largely attributed to three major factors. First, the adaptive feature selection component eliminates redundant and irrelevant traffic features to concentrate on only discriminative characteristics. Second, the CNN-BiLSTM model architecture can simultaneously identify local and long-range traffic communication patterns. Third, the attention component dynamically assigns importance to cybersecurity-relevant attributes for improved attack detection capability and enhanced generalization performance. In conclusion, the comparative analysis demonstrates that the proposed CNN-BiLSTM-Attention framework outperforms the state-of-the-art machine learning and deep learning frameworks. The empirical analysis shows that the developed CNN-BiLSTM-Attention framework can achieve accurate cyberattack detection without sacrificing classification robustness when applied to various datasets. The empirical findings have provided a firm basis for subsequent resilient-based analysis, which includes ablation, dynamic risk analysis, automated recovery, and cyber resilience analyses presented in the upcoming sections.

5 Conclusion

In this study, an AI-Enhanced Cyber Resilience Framework for critical infrastructure protection has been introduced, incorporating adaptive feature selection, hybrid deep learning-based cyber threat detection, risk assessment in real-time, and automated recovery optimization into a common framework. In the process, MI-based feature selection technique and CNN-BiLSTM-Attention based cyber threat detection algorithm have been used, and TIS, NLI, and APS parameters have been considered for attack characterization. Experimentation conducted using UNSW-NB15 and CICIDS2017 datasets shows that the proposed cyber resilience framework is capable of delivering improved results in terms of threat detection accuracy, reduced mean time to recovery, enhanced recovery success rate, and maintaining high availability under various types of cyberattacks.

5.1 Future Work

Research in the future will center around the validation of the proposed approach within real-world operations along with the application of some sophisticated techniques involving Artificial Intelligence like Graph Neural Networks (GNNs), Transformers, and reinforcement learning. Moreover, the framework can be adapted to new areas that emerge with respect to cyber-physical domains like smart grids, Industry 5.0, intelligent transport networks, and smart healthcare systems.

Data Availability Statement

The data sources that were used in the current research are public. In this respect, the UNSW-NB15 dataset is publicly available through the Australian Centre for Cyber Security (ACCS) and Kaggle platforms, whereas the CICIDS2017 dataset was provided publicly by the Canadian Institute for Cybersecurity (CIC). All the datasets that have been used in this research are freely available. The preprocessed data sets and experiment settings performed to generate results in this research can be provided upon request to the corresponding author.
UNSW-NB15 Dataset: https://research.unsw.edu.au/projects/unsw-nb15-dataset

UNSW-NB15 (Kaggle): https://www.kaggle.com/datasets/mrwellsdavid/unsw-nb15
CICIDS2017 Dataset: https://www.unb.ca/cic/datasets/ids-2017.html

References

[1] inkov, I., & Trump, B. D. (2019). *The science and practice of resilience* (pp. 110-115). Cham: Springer International Publishing.

[2] ational Institute of Standards and Technology (US), & National Institute of Standards and Technology (US). (2024). *NIST Cybersecurity Framework 2.0: Quick-start Guide for Creating and Using Organizational Profiles*. US Department of Commerce, National Institute of Standards and Technology.

[3] . D. Trump, M. Cegan, and I. Linkov, "Resilience and risk management in critical infrastructure systems," *Environment Systems and Decisions*, 2022.

[4] ct, C. R. (2024). Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements. *Official Journal of the European Union, L series*, 10.

[5] iafe, I., Koranteng, F. N., Obeng, E. N., Assyne, N., Wiafe, A., & Gulliver, S. R. (2020). Artificial intelligence for cybersecurity: a systematic mapping of literature. *IEEE Access*, 8, 146598-146612.

[6] leesa, A. M., Zaidan, B. B., Zaidan, A. A., & Sahar, N. M. (2020). Review of intrusion detection systems based on deep learning techniques: coherent taxonomy, challenges, motivations, recommendations, substantial analysis and future directions. *Neural Computing and Applications*, 32(14), 9827-9858.

[7] o Xuan, C., Dao, M. H., & Nguyen, H. D. (2020). APT attack detection based on flow network analysis techniques using deep learning. *Journal of Intelligent & Fuzzy Systems*, 39(3), 4785-4801.

[8] hester, M. V., & Allenby, B. R. (2020). Perspective: the cyber frontier and infrastructure. *IEEE Access*, 8, 28301-28310.

[9] lowononi, F. O., Rawat, D. B., & Liu, C. (2020). Resilient machine learning for networked cyber physical systems: A survey for machine learning security to securing machine learning for CPS. *IEEE Communications Surveys & Tutorials*, 23(1), 524-552.

[10] lmaiah, M. A., Saqr, L. M., Al-Rawwash, L. A., Altellawi, L. A., Al-Ali, R., & Almomani, O. (2024). Classification of Cybersecurity Threats, Vulnerabilities and Countermeasures in Database Systems. *Computers, Materials & Continua*, 81(2).

[11] alvi, A., Spagnoletti, P., & Noori, N. S. (2022). Cyber-resilience of critical cyber infrastructures: Integrating digital twins in the electric power ecosystem. *Computers & Security*, 112, 102507.

[12] lkhaleel, B. A. (2024). Machine learning applications in the resilience of interdependent critical infrastructure systems—A systematic literature review. *International Journal of Critical Infrastructure Protection*, 44, 100646.

[13] . Järveläinen, J. Heikkilä, and M. Siponen, "Towards a framework for improving cybersecurity resilience," *Enterprise Information Systems*, 2025, doi: 10.1080/12460125.2025.2479546.

[14] . Khdhir, "Building resilient national critical infrastructure: A digital twin approach," *Discover Applied Sciences*, 2026, doi: 10.1007/s44443-026-00464-5.

[15] akke, D. G. (2025). PERFORMANCE EVALUATION OF MACHINE LEARNING-BASED INTRUSION DETECTION USING NSL-KDD, UNSW-NB15 AND CICIDS2017 DATASETS. *International Journal of Applied Mathematics*, 38(3s), 447-469.

[16] . More, M. Idrissi, H. Mahmoud, and A. T. Asyhari, "Enhanced intrusion detection systems performance with UNSW-NB15 data analysis," *Algorithms*, vol. 17, no. 2, p. 64, 2024, doi: 10.3390/a17020064.

[17] . A. Ajagbe, A. A. Adegun, A. B. Olanrewaju, J. B. Oladosu, and M. O. Adigun, "Intrusion detection: A comparison study of machine learning algorithms," *SN Computer Science*, vol. 5, 2024, doi: 10.1007/s42979-024-03369-0.

[18] . Arcos-Argudo, A. García-Sánchez, and J. Martínez, "A deterministic comparison of classical machine learning methods for network intrusion detection," *Algorithms*, vol. 18, no. 12, p. 749, 2025, doi: 10.3390/a18120749.

[19] . Fatani, M. Abd Elaziz, A. Dahou, M. A. A. Al-Qaness, and S. Lu, "Enhancing intrusion detection systems for IoT and cloud environments using deep learning and optimization," *Sensors*, vol. 23, no. 9, p. 4430, 2023, doi: 10.3390/s23094430.

[20] . Bovenzi, G. Aceto, D. Ciuonzo, A. Montieri, V. Persico, and A. Pescapé, "Network anomaly detection methods in IoT environments via deep learning: A fair comparison of performance and robustness," *Computers & Security*, vol. 128, p. 103167, 2023, doi: 10.1016/j.cose.2023.103167.

[21] . T. Azar, E. Shehab, A. M. Mattar, I. A. Hameed, and S. A. Elsaid, "Deep learning based hybrid intrusion detection systems to protect satellite networks," *Journal of Network and Systems Management*, vol. 31, p. 82, 2023, doi: 10.1007/s10922-023-09767-8.

[22] . Awajan, "A novel deep learning-based intrusion detection system for IoT networks," *Computers*, vol. 12, no. 2, p. 34, 2023, doi: 10.3390/computers12020034.

[23] . Gombar and A. Topalović, "Cost-aware lightweight deep learning for intrusion detection: A comparative study on UNSW-NB15 and CIC-IDS2017," *Electronics*, vol. 15, no. 8, p. 1603, 2026, doi: 10.3390/electronics15081603.

[24] . Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," *ICISSP*, 2018, doi: 10.5220/0006639801080116.

[25] . Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," *MilCIS*, 2015, doi: 10.1109/MilCIS.2015.7348942.

[26] . Dube, "Faulty use of the CIC-IDS 2017 dataset in information security research," *Journal of Computer Virology and Hacking Techniques*, vol. 20, pp. 203–211, 2024, doi: 10.1007/s11416-023-00509-7.

[27] . Wang, X. Zhang, and Y. Li, "Learn-IDS: Bridging gaps between datasets and real-world intrusion detection deployment," *Electronics*, vol. 13, no. 6, p. 1072, 2024, doi: 10.3390/electronics13061072.

[28] . Roy, J. Li, B.-J. Choi, and Y. Bai, "A lightweight supervised intrusion detection mechanism for IoT networks," *Future Generation Computer Systems*, vol. 127, pp. 276–285, 2022, doi: 10.1016/j.future.2021.09.027.

[29] . Lahasan and H. Samma, "Optimized deep auto-encoder model for Internet of Things intruder detection," *IEEE Access*, vol. 10, pp. 8434–8448, 2022, doi: 10.1109/ACCESS.2022.3144208.

[30] . Saba, A. Rehman, T. Sadad, H. Kolivand, and S. A. Bahaj, "Anomaly-based intrusion detection system for IoT networks through deep learning model," *Computers and Electrical Engineering*, vol. 99, p. 107810, 2022, doi: 10.1016/j.compeleceng.2022.107810.

[31] . Zhang and Y. Zhang, "Intrusion detection model for industrial Internet of Things based on improved auto-encoder," *Computational Intelligence and Neuroscience*, 2022, doi: 10.1155/2022/1406214.

[32] . Marzouk et al., "Hybrid deep learning enabled intrusion detection in clustered IIoT environment," *Computers, Materials & Continua*, vol. 72, no. 2, pp. 3763–3775, 2022, doi: 10.32604/cmc.2022.027483.

[33] . Kumar, R. Singh, and P. K. Sharma, "Impact of machine learning on intrusion detection systems for critical infrastructure," *Information*, vol. 16, no. 7, p. 515, 2025, doi: 10.3390/info16070515.

[34] . Fatani, M. Abd Elaziz, A. Dahou, M. A. A. Al-Qaness, and S. Lu, "IoT intrusion detection system using deep learning and enhanced transient search optimization," *IEEE Access*, vol. 9, pp. 123448–123464, 2021, doi: 10.1109/ACCESS.2021.3109081.

[35] . Wu, L. Nie, S. Wang, Z. Ning, and S. Li, "Intelligent intrusion detection for Internet of Things security: A deep convolutional generative adversarial network-enabled approach," *IEEE Internet of Things Journal*, vol. 10, no. 4, pp. 3094–3106, 2023, doi: 10.1109/JIOT.2021.3112159.

[36] ankotia, S., Conte de Leon, D., & Rimal, B. P. (2026). FedPrIDS: Privacy-Preserving Federated Learning for Collaborative Network Intrusion Detection in IoT. *Journal of Cybersecurity and Privacy*, 6(1), 10.