Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

REAL-TIME EDGE AI FOR SECURE SMART GRID IOT–CLOUD ECOSYSTEMS: AN INTELLIGENT EDGE COMPUTING FRAMEWORK FOR CYBERATTACK DETECTION, PRIVACY-PRESERVING ENERGY ANALYTICS AND ADAPTIVE THREAT RESPONSE

Muhammad Nauman Hameed[1], Muhammad Sajjad[*2], Tanveer Ahmad[3], Shahzeb Iqbal[4], Junaid Ali[5]

[1]*Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan.*
[*2]*Chief Executive Officer, Gitchia Institute of Global Certification, Lahore, Pakistan.*
[3]*Department of Cybersecurity, New York University, New York, USA.*
[4]*Department of Computer Science, University of Mianwali, Mianwali, Pakistan.*
[5]*Department of Computer and Information Engineering, Islamic International University, Malaysia.*

[1]naumanhameed149@gmail.com, [*2]mrsajjad201@yahoo.com, [3]tanveerahmad66@gmail.com, [4]shahzebiqbal749@gmail.com, [5]junaidali445@gmail.com

DOI: https://doi.org/10.5281/zenodo.22659404

Keywords
Smart grid; Edge AI; Internet of Things; cybersecurity; cyberattack detection; federated learning; privacy-preserving analytics; adaptive threat response.

Article History
Received: 30 January 2026
Accepted: 15 March 2026
Published: 31 March 2026

Corresponding Author: *
Muhammad Sajjad

Abstract

The rapid digitalization of smart grids through Internet of Things (IoT) devices, advanced metering infrastructure, distributed energy resources, and cloud-based energy management has improved operational efficiency while simultaneously expanding the cyberattack surface. Conventional cloud-centered security mechanisms often experience communication delays, privacy risks, bandwidth limitations, and inadequate responsiveness to rapidly evolving threats. This study presents a real-time Edge Artificial Intelligence framework for securing smart-grid IoT–cloud ecosystems through distributed cyberattack detection, privacy-preserving energy analytics, and adaptive threat response. A multi-source dataset containing 285,000 operational and cybersecurity records was constructed from smart meters, phasor measurement units, substation sensors, network traffic, authentication logs, and energy-consumption profiles. The dataset incorporated normal activity and six major attack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. Following data cleaning, temporal synchronization, normalization, class balancing, and feature selection, 48 predictive variables were retained. The proposed architecture combines a lightweight convolutional neural network, bidirectional long short-term memory layers, and an attention mechanism to capture spatial, temporal, and context-sensitive attack patterns at the network edge. Federated learning and differential privacy were incorporated to support collaborative model training without transmitting raw consumer or grid-operational data to the cloud. Experimental evaluation used stratified training, validation, and testing partitions of 70%, 15%, and 15%, respectively. The proposed model achieved 98.1% accuracy, 97.8% precision, 97.6% recall, a 97.7% F1-score, and a receiver operating characteristic area under the curve of

0.993. It outperformed logistic regression, support vector machine, random forest, and conventional CNN models, which achieved accuracies of 84.7%, 89.6%, 93.8%, and 95.4%, respectively. Edge-based inference reduced average threat-detection latency from 186 milliseconds in the cloud-only configuration to 42 milliseconds, representing a 77.4% reduction, while decreasing cloud-bound data transmission by 68.2%. The adaptive response mechanism isolated compromised devices and prioritized security alerts, reducing successful attack propagation by 71.6%. Differentially private federated learning preserved sensitive energy data with only a 1.2-percentage-point reduction in detection accuracy. These findings demonstrate that the framework provides accurate, low-latency, privacy-aware, and scalable cybersecurity for next-generation smart-grid infrastructure.

1- INTRODUCTION:

The rapid digital transformation of electrical power systems has accelerated the development of smart grids characterized by bidirectional energy flow, decentralized generation, automated control, and real-time communication. Smart meters, phasor measurement units (PMUs), intelligent electronic devices, substation sensors, distributed energy resources, electric vehicles, and cloud-based energy-management platforms now continuously exchange large volumes of operational data. These technologies improve grid observability, demand forecasting, fault management, resource allocation, and consumer participation. However, connecting geographically distributed physical infrastructure to Internet of Things (IoT) networks and cloud platforms has also expanded the grid's cyberattack surface. A successful intrusion can manipulate measurements, interrupt communication, compromise consumer privacy, destabilize control processes, and potentially cause large-scale power disruption. Smart-grid IoT–cloud ecosystems are particularly vulnerable because they combine heterogeneous devices, communication protocols, computing environments, and security requirements. Many field devices possess limited processing power, memory, and energy capacity, making the deployment of conventional security mechanisms difficult [1]. Legacy equipment may also lack secure authentication, encrypted communication, and reliable software-update capabilities. Moreover, operational technology networks increasingly interact with information technology and cloud services, creating additional entry points for adversaries. False-data injection attacks can alter sensor measurements and mislead state-estimation systems, while denial-of-service attacks can disrupt communication between grid components. Energy-theft attacks manipulate consumption records, command-injection attacks interfere with control operations, device-spoofing attacks impersonate legitimate assets, and malware-based intrusions enable persistent unauthorized access. The diversity and evolving nature of these attacks require intelligent security mechanisms capable of identifying complex spatial and temporal patterns. Traditional intrusion-detection systems commonly depend on predefined rules, static signatures, or centralized cloud processing. Signature-based approaches can effectively identify previously observed attacks but frequently fail against new, modified, or zero-day threats. Rule-based mechanisms can also generate excessive false alarms when legitimate grid behavior changes because of variations in demand, weather, distributed generation, or device operating conditions [2]. Centralized machine-learning solutions offer improved analytical capability but require large volumes of smart-grid data to be transmitted to remote cloud servers. This arrangement introduces communication latency, increases bandwidth consumption, creates potential single points of failure, and exposes sensitive consumer and operational information during transmission and storage. Such limitations are especially critical in smart-grid environments, where delayed security decisions may allow an intrusion to propagate through interconnected devices before a response is initiated. Edge computing provides an effective alternative by

moving data processing and intelligent decision-making closer to smart meters, substations, gateways, and other data sources. Instead of transmitting every observation to the cloud, edge nodes can preprocess data, execute trained detection models, and generate security decisions locally. This distributed approach can reduce network congestion and detection latency while maintaining essential cybersecurity functions during intermittent cloud connectivity. Edge-based intelligence also supports context-aware monitoring because local nodes can analyze device behavior and operational conditions in near real time [3]. Nevertheless, the resource limitations of edge devices require lightweight artificial-intelligence models that balance detection performance with computational efficiency. A model that is highly accurate but too complex for real-time edge deployment may provide limited practical value.

Deep learning offers promising capabilities for detecting cyberattacks in complex smart-grid data. Convolutional neural networks (CNNs) can identify local correlations and discriminative patterns among multivariate measurements, communication statistics, authentication events, and device-level features. Bidirectional long short-term memory (BiLSTM) networks can analyze temporal dependencies in both forward and backward directions, enabling the detection of abnormal sequences that may remain hidden in individual observations. Attention mechanisms can further improve detection by assigning greater importance to the most relevant time steps and features. The integration of these components can therefore capture spatial, temporal, and contextual characteristics of smart-grid attacks. However, conventional deep-learning systems are often centrally trained and may require the collection of raw information from multiple consumers, substations, and grid operators. Energy data contain sensitive information about consumer activities, occupancy patterns, appliance usage, industrial operations, and grid conditions. Transmitting such data to a central location can create significant privacy and regulatory concerns. Federated learning addresses this limitation by allowing participating edge nodes to train a shared model without directly transferring their raw datasets [4]. Each node performs local training and communicates model updates to an aggregation server, which combines the updates to construct an improved global model. Although this approach reduces direct data exposure, information may still be inferred from shared gradients or model parameters. Differential privacy can mitigate this risk by introducing carefully controlled statistical noise into model updates, thereby limiting the disclosure of individual records. Combining federated learning with differential privacy can consequently support collaborative cyberattack detection while protecting consumer and infrastructure data. Detection alone is insufficient for protecting a highly interconnected smart-grid environment. Security systems must also determine how to respond according to attack type, confidence level, affected asset, and operational risk. A conventional system that merely forwards every alert to a central security team may produce delays and alert overload. An adaptive response mechanism can prioritize high-risk incidents and automatically execute predefined protective actions, including isolating compromised devices, restricting suspicious network traffic, revoking authentication tokens, switching to trusted control channels, and requesting human validation. Such responses must remain context-sensitive because an unnecessary device disconnection may affect grid availability. Accordingly, automated actions should be coordinated with confidence thresholds, asset criticality, and human-in-the-loop oversight. Despite recent progress in edge computing and artificial intelligence, several research limitations remain. Many existing studies examine attack detection, edge inference, privacy protection, or automated response as separate problems. Some evaluate models using limited datasets or a narrow set of attacks, reducing their ability to represent the heterogeneity of practical smart-grid environments. Other studies emphasize classification accuracy without assessing detection latency, communication overhead, privacy–utility trade-offs, or attack propagation. Moreover, comparatively limited research has developed an

integrated framework that combines lightweight deep learning, distributed privacy-preserving training, and adaptive response across interconnected smart-grid IoT and cloud layers [5]. These gaps demonstrate the need for a unified solution that considers detection performance, computational efficiency, data privacy, scalability, and operational responsiveness simultaneously. To address these challenges, this study proposes a real-time Edge AI framework for secure smart-grid IoT–cloud ecosystems. The framework integrates a lightweight CNN, BiLSTM layers, and an attention mechanism to detect false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. A multi-source dataset containing 285,000 operational and cybersecurity records is constructed from smart meters, PMUs, substation sensors, network traffic, authentication logs, and energy-consumption profiles. After data cleaning, temporal synchronization, normalization, class balancing, and feature selection, 48 predictive variables are retained. The trained detection model is deployed at the edge to enable rapid inference, while federated learning and differential privacy support secure collaborative training without centralizing raw data. An adaptive response component then converts detection outcomes into risk-prioritized security actions.

The principal contributions of this study are as follows:

1. A unified Edge AI architecture is developed for real-time cyberattack detection across heterogeneous smart-grid IoT, edge, and cloud environments.
2. A lightweight CNN–BiLSTM–Attention model is introduced to jointly capture spatial feature relationships, bidirectional temporal dependencies, and context-sensitive attack indicators.
3. Six important attack categories false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion are evaluated using a large multi-source dataset containing 285,000 records and 48 selected variables.
4. Federated learning and differential privacy are integrated to enable collaborative security-model development while limiting the exposure of consumer energy data and sensitive grid-operational information.
5. An adaptive threat-response mechanism is designed to prioritize alerts, isolate compromised devices, and restrict attack propagation according to detection confidence and asset risk.
6. The framework is evaluated using classification performance, inference latency, cloud-bound data transmission, privacy-related accuracy loss, and attack-propagation reduction, providing a multidimensional assessment of its operational effectiveness.

The experimental findings demonstrate that the proposed framework achieves 98.1% detection accuracy and an area under the receiver operating characteristic curve of 0.993. Edge-based inference reduces average detection latency from 186 ms to 42 ms and decreases cloud-bound data transmission by 68.2%. The adaptive response mechanism reduces successful attack propagation by 71.6%, while differentially private federated learning limits the accuracy reduction to 1.2 percentage points. These results indicate that the proposed framework can provide accurate, low-latency, privacy-aware, and scalable cybersecurity for next-generation smart-grid infrastructure.

2- Deep Neural Intelligence for Spatial–Temporal Attack Discovery:

Deep-learning techniques have become increasingly important for detecting sophisticated cyberattacks in smart-grid IoT–cloud ecosystems. Unlike conventional machine-learning algorithms, which often require manually constructed variables, deep neural networks can automatically learn hierarchical representations from high-dimensional operational and cybersecurity data. These models can jointly analyse smart-meter measurements, PMU signals, substation conditions, network traffic, authentication events, and energy-consumption patterns. This capability is especially valuable for detecting attacks that produce small but coordinated changes across multiple data sources. Convolutional neural networks are widely used to

extract local and spatial relationships from multivariate smart-grid records [6]. Their convolutional filters can identify recurring patterns involving voltage deviations, abnormal packet rates, inconsistent device identities, unexpected command sequences, and irregular energy consumption. CNNs are computationally efficient because of parameter sharing and local connectivity, making lightweight versions suitable for deployment on edge gateways. However, conventional CNNs mainly emphasize local feature patterns and may not sufficiently capture the long-term progression of coordinated or multistage cyberattacks. Recurrent neural networks address this limitation by examining sequential dependencies among observations. In particular, long short-term memory networks use internal memory gates to retain important information across extended sequences. They can therefore detect gradual attack development, repeated authentication failures, abnormal communication intervals, and persistent manipulation of grid measurements. Bidirectional LSTM networks further strengthen temporal analysis by processing each sequence in forward and backward directions. This enables the model to evaluate an event according to both its preceding activity and subsequent context. The comparative characteristics of the principal deep-learning components are summarized in Table 1.

Table 1: Comparative roles of deep-learning components in smart-grid cyberattack detection

Deep-learning component	Primary analytical role	Smart-grid patterns captured
CNN	Spatial and local feature extraction	Measurement deviations, packet patterns and feature interactions
LSTM	Sequential dependency analysis	Gradual anomalies, repeated access attempts and attack progression
BiLSTM	Bidirectional temporal analysis	Preceding and subsequent event relationships
Attention mechanism	Context-sensitive feature prioritization	Critical time steps, variables and abnormal events
CNN–BiLSTM–Attention	Integrated spatial–temporal detection	Local, sequential and contextual attack signatures

Attention mechanisms have recently been integrated with recurrent architectures to improve context-sensitive threat recognition. Instead of treating every feature and time step as equally important, attention assigns greater weight to the hidden representations that contribute most strongly to a classification decision. For example, the model may emphasize a sudden phase-angle change during a false-data injection attack, an abnormal packet-rate increase during a denial-of-service event, or repeated identity inconsistencies during device spoofing. This selective focus can improve detection accuracy while also providing operators with information about the events that influenced the prediction [7]. The integration of CNN, BiLSTM, and attention layers provides a comprehensive approach to spatial–temporal attack discovery. As illustrated in Figure 1, raw smart-grid records are first arranged into synchronized sequences. The CNN extracts short-range correlations and local attack indicators, after which the BiLSTM examines their temporal progression in both directions. The attention mechanism then prioritizes the most relevant representations before the final classification layer distinguishes normal activity from the six attack categories.

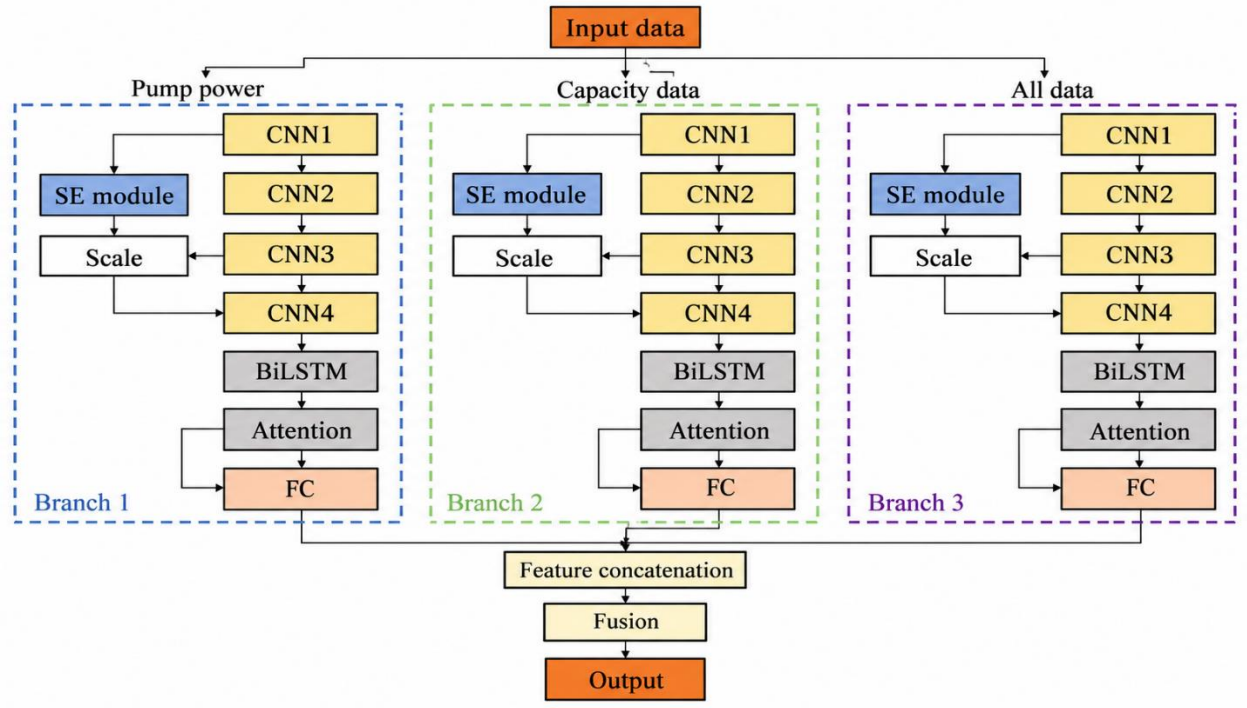

Figure 1: Three-branch CNN–BiLSTM–attention architecture for multi-source feature extraction and fusion.

This hybrid structure is particularly appropriate for identifying false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. Each attack produces a different combination of electrical, behavioural, network, and temporal indicators. A CNN-only model may identify local anomalies but overlook their progression, whereas a recurrent model may process temporal relationships without extracting sufficiently detailed feature interactions. The combined architecture addresses these limitations by learning local patterns, sequential behaviour, and contextual importance within a unified framework. Nevertheless, deploying deep neural networks at the edge introduces challenges related to memory, processing capacity, energy consumption, and inference latency [8]. Consequently, the architecture must be optimized through lightweight convolutional filters, controlled BiLSTM dimensions, dropout regularization, and efficient sequence processing. Such optimization enables high-quality attack discovery without exceeding the limited computational capacity of smart-grid edge devices. Therefore, spatial–temporal deep learning offers a strong analytical foundation for accurate, low-latency, and context-aware protection of distributed smart-grid infrastructure.

3- Attention Mechanisms for Context-Aware Threat Recognition:

Attention mechanisms have emerged as an effective solution for improving context-aware cyberattack detection in smart-grid IoT–cloud ecosystems. Smart grids continuously produce large volumes of heterogeneous information from smart meters, phasor measurement units, substation sensors, network traffic, authentication systems, control devices, and energy-consumption profiles. Although these data sources provide valuable security information, not every variable or time interval contributes equally to attack detection. Attention mechanisms address this issue by enabling deep-learning models to prioritize the observations most relevant to a particular cybersecurity event. Conventional sequential models generally compress an entire

input sequence into a limited internal representation. During this process, subtle attack indicators may lose their influence, especially when a long sequence contains a large proportion of normal activity. Attention-based models overcome this limitation by evaluating the relative importance of the representations generated at different time steps. Greater importance is assigned to unusual measurements, suspicious communication behaviour, identity inconsistencies, unauthorized commands, and other indicators that strongly support an attack classification [9]. The security indicators prioritized by the attention mechanism vary according to the nature of the attack. During a false-data injection attack, the model may focus on inconsistencies among voltage, frequency, phase angle, and expected grid states. In denial-of-service detection, sudden increases in packet volume, repeated service requests, connection failures, and communication delays may receive greater attention. For energy theft, the model can emphasize deviations between current and historical consumption profiles. Similarly, unauthorized control instructions become important for command-injection detection, whereas identity conflicts and irregular authentication patterns support the recognition of device-spoofing attacks. The principal contextual indicators associated with the six attack categories are presented in Table 2.

Table 2: Context-sensitive indicators prioritized during attention-based threat recognition

Attack category	Indicators receiving greater attention	Contextual relationship examined
False-data injection	Voltage deviations, phase-angle changes and inconsistent sensor readings	Consistency between reported measurements and expected grid states
Denial-of-service	Packet bursts, request frequency, connection failures and delays	Sudden changes in communication behaviour
Energy theft	Consumption deviations, meter-state changes and unusual load reductions	Difference between current and historical energy use
Command injection	Unauthorized instructions, control-message frequency and unexpected device responses	Agreement between user privileges and issued commands
Device spoofing	Identity conflicts, access-origin changes and authentication anomalies	Consistency among device identity, location and network behaviour
Malware-based intrusion	Abnormal processes, repeated connections and unusual data transfers	Combined device, network and behavioural activity

Within the proposed architecture, CNN and BiLSTM layers first generate spatial–temporal representations from synchronized smart-grid data. The attention layer then examines these representations and assigns greater importance to the elements that provide the strongest evidence of suspicious behaviour. The selected information is combined into a context-rich representation and forwarded to the classification layer. The model subsequently produces the predicted threat category and its associated confidence level. This process is illustrated in Figure 2.

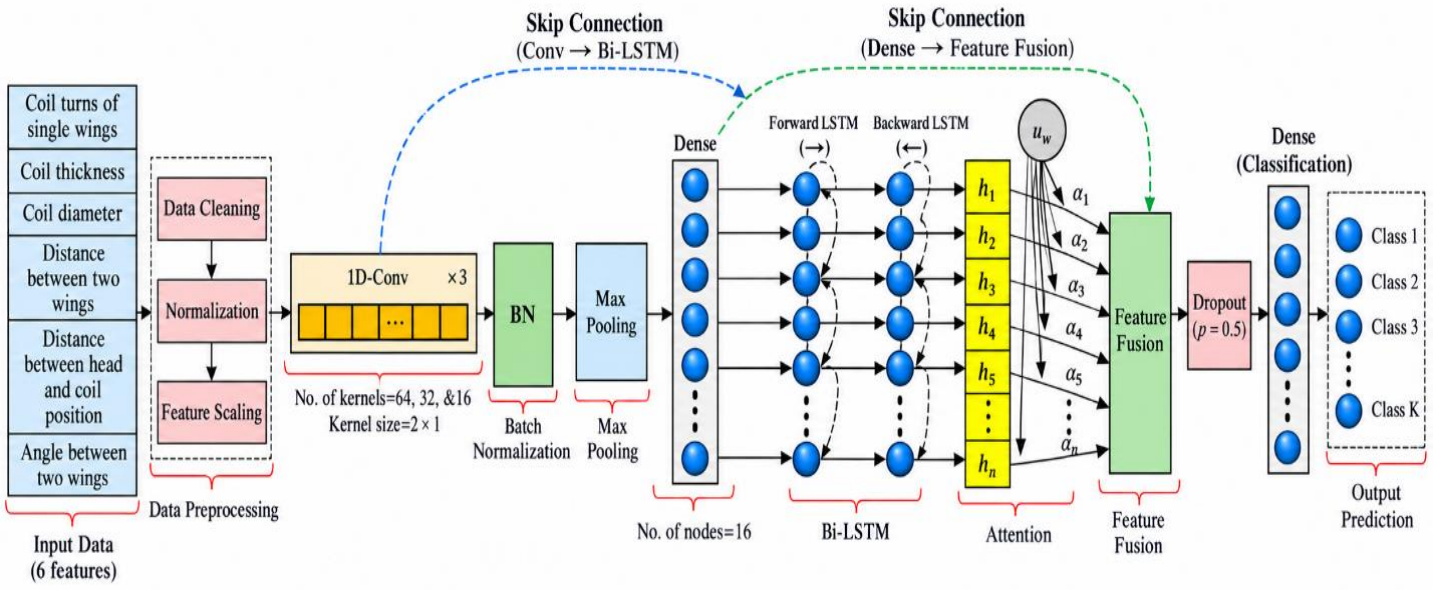

Figure 2: Attention-based process for context-aware smart-grid threat recognition.

The integration of attention with CNN–BiLSTM processing provides several advantages. First, it reduces the influence of irrelevant or repetitive observations that may otherwise weaken the classification decision. Second, it enables the model to distinguish cyberattacks from legitimate operational changes caused by variations in demand, renewable-energy generation, weather conditions, or switching activities. Third, it improves the recognition of multistage attacks by identifying critical events distributed across a sequence rather than relying on a single observation. Attention information can also support security analysts and grid operators [10]. When an alert is generated, the system can highlight the variables and time intervals that contributed most strongly to the prediction. For example, a command-injection alert may be associated with an unauthorized instruction, an unusual control-message frequency, and an unexpected device response. Presenting these indicators can assist with alert validation, forensic investigation, and human-in-the-loop decision-making. Nevertheless, attention weights represent the model's internal focus and should be interpreted as supporting evidence rather than complete causal explanations. Despite these benefits, attention mechanisms introduce additional computational operations. An unnecessarily complex attention layer may increase memory requirements and inference time, reducing its suitability for resource-constrained smart-grid edge devices [11]. Therefore, lightweight attention components, controlled input-sequence lengths, and optimized model dimensions are required for practical implementation. When properly configured, attention-based processing strengthens contextual threat recognition, improves multi-class attack discrimination, and provides valuable evidence for adaptive security responses across distributed smart-grid IoT–cloud environments.

4- Methodology:

This study employed a quantitative experimental methodology to develop and evaluate a real-time Edge AI framework for securing smart-grid IoT–cloud ecosystems. The research process integrated multi-source data acquisition, data preprocessing, cyberattack labelling, feature engineering, deep-learning model development, privacy-preserving federated training, edge-based inference, and adaptive threat response. A lightweight CNN–BiLSTM–Attention architecture was designed to

capture local feature interactions, bidirectional temporal relationships, and context-sensitive indicators of malicious behaviour. The framework was developed to classify normal grid activity and six major attack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. The experimental evaluation examined both predictive and operational performance [12]. The dataset was divided into training, validation, and testing subsets using a stratified 70:15:15 ratio. Logistic regression, support vector machine, random forest, and conventional CNN models were used as comparative baselines. In addition to accuracy, precision, recall, F1-score, and ROC–AUC, the framework was assessed according to detection latency, cloud-bound data transmission, privacy-related accuracy loss, and reduction in successful attack propagation. Federated learning and differential privacy were incorporated to support collaborative model development without transferring raw consumer or grid-operational data to the cloud.

4.1- Multi-Source Smart-Grid Dataset Construction:

A multi-source dataset containing 285,000 operational and cybersecurity records was constructed to represent the heterogeneous behaviour of smart-grid IoT–cloud ecosystems. The dataset integrated information from smart meters, phasor measurement units (PMUs), substation sensors, network traffic monitors, authentication logs, and energy-consumption profiles. Combining these sources enabled the framework to examine electrical measurements, equipment conditions, communication behaviour, access activities, and consumer-energy patterns within a unified analytical environment. Smart-meter records included voltage, current, active and reactive power, power factor, meter status, reporting intervals, and consumption deviations. PMU data contained voltage and current phasors, system frequency, phase angle, and rate-of-change-of-frequency measurements [13]. These electrical variables provided important evidence for detecting false-data injection and identifying inconsistencies between reported measurements and actual grid conditions. Substation-sensor data included transformer loading, equipment temperature, relay activity, feeder status, breaker conditions, and protection-system events. These records supplied the operational context required to distinguish malicious activity from legitimate switching operations or equipment faults. Network traffic data included packet frequency, protocol type, connection duration, payload size, communication delay, retransmission activity, and failed connection attempts. Such variables supported the identification of denial-of-service attacks, malware communication, unauthorized scanning, and unusual device interactions. Authentication logs recorded successful and failed login attempts, credential status, access origin, device identity, authorization level, session duration, and privilege changes [14]. Energy-consumption profiles contained current and historical demand, peak usage, short-term load variation, consumption regularity, and differences between expected and observed energy use. The characteristics and analytical roles of the six data sources are summarized in Table 3.

Table 3: Data sources used in constructing the smart-grid cybersecurity dataset

Data source	Operational purpose	Relevant attack evidence
Smart meters	Consumer-level energy monitoring	Energy theft and false-data injection
PMUs	High-resolution grid-state observation	Measurement manipulation and coordinated grid attacks
Substation sensors	Equipment and substation monitoring	Command injection and malware-related disruption
Network traffic monitors	Communication-network monitoring	Denial-of-service and malware intrusion

Authentication logs	Identity and access management	Device spoofing and unauthorized access
Energy-consumption profiles	Behavioural and demand-pattern analysis	Energy theft and consumption manipulation

Each record was assigned a standardized timestamp, anonymized device identifier, source category, operational location, communication status, and security label. Because the data sources operated at different sampling frequencies, temporal synchronization was applied to associate related electrical, network, authentication, and equipment events. For example, an unusual control instruction was linked with the corresponding login event, device response, and substation condition. This integration improved the ability to identify coordinated attacks that might remain undetected when each data source is examined independently. The dataset represented normal activity and six major attack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion [15]. Labels were assigned using attack signatures, network-event evidence, authentication outcomes, device-state changes, and deviations from expected grid behaviour. Observations with insufficient or conflicting evidence were excluded to maintain labelling reliability. The dataset-construction process is illustrated in Figure 3. Heterogeneous records were first collected from the six smart-grid sources and then passed through identity anonymization, timestamp standardization, temporal integration, and security labelling. The integrated dataset was subsequently prepared for cleaning, feature engineering, and model development.

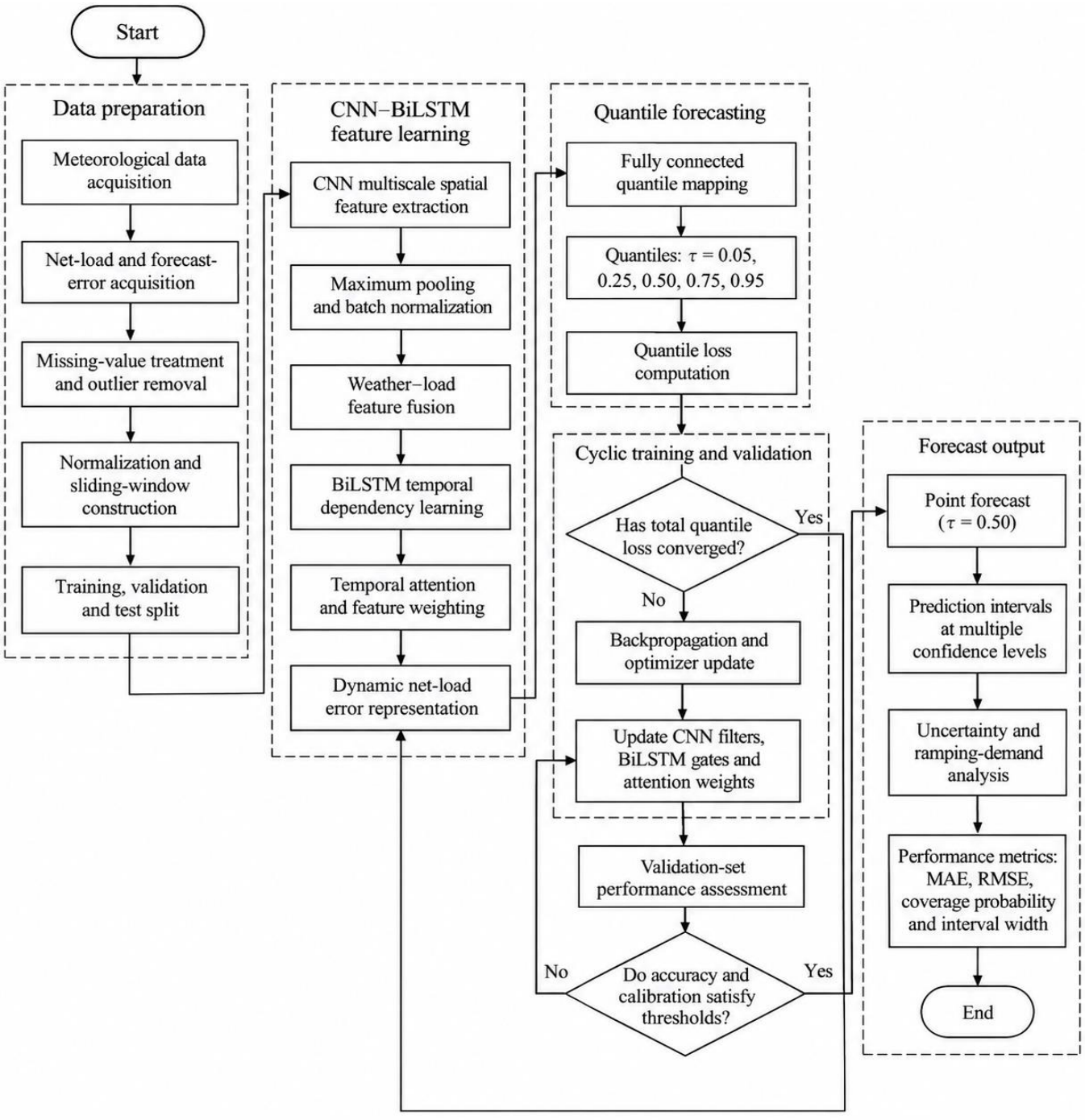

Figure 3: Construction and integration of the multi-source smart-grid cybersecurity dataset.

Following integration, the records underwent duplicate removal, missing-value treatment, measurement standardization, class balancing, and feature selection. Redundant and low-information variables were removed, while variables with strong operational or cybersecurity relevance were retained [16]. The final dataset contained 48 predictive variables for training and evaluating the proposed CNN–BiLSTM–Attention framework. This multi-source

construction improved the dataset's ability to represent real-world smart-grid conditions and supported the detection of spatial, temporal, behavioural, and context-dependent cyberattack patterns.

4.2- Threat Library and Multi-Class Attack Labelling:

A structured threat library was developed to support consistent identification and multi-class labelling of cybersecurity events within the smart-grid dataset. The library represented normal grid activity and six major attack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. These categories were selected because they affect different security objectives, including data integrity, service availability, consumer privacy, device authenticity, access control, and operational safety. False-data injection records represented unauthorized manipulation of smart-meter, PMU, or substation measurements. These attacks included abnormal changes in voltage, frequency, phase angle, power flow, and energy readings that were inconsistent with related grid conditions [17]. Denial-of-service records represented packet flooding, repeated service requests, communication congestion, connection failures, and unusual delays intended to reduce the availability of grid services. Energy-theft labels identified manipulation of meter readings, consumption records, billing information, or device configurations. Command-injection attacks involved unauthorized control instructions directed toward breakers, relays, gateways, or distributed energy resources. Device-spoofing labels represented cases in which a malicious or compromised node impersonated an authorized smart-grid device [18]. Malware-based intrusion included abnormal process execution, unauthorized data transfer, persistent external communication, privilege escalation, and lateral movement across connected infrastructure. The threat definitions and principal labelling indicators are summarized in Table 4.

Table 4: Smart-grid threat library and multi-class labelling criteria

Label code	Event category	Potential operational consequence
C0	Normal activity	Stable grid operation
C1	False-data injection	Incorrect state estimation and control decisions
C2	Denial-of-service	Communication or service interruption
C3	Energy theft	Revenue loss and inaccurate demand estimation
C4	Command injection	Improper switching or equipment operation
C5	Device spoofing	Unauthorized network participation
C6	Malware-based intrusion	Device compromise and lateral attack propagation

Each record was assigned one of the seven labels using evidence from multiple data sources. The labelling process considered attack signatures, network behaviour, authentication outcomes, device-state changes, energy-profile deviations, and temporal relationships among connected events. For instance, an unusual breaker operation was not automatically classified as command injection. It was labelled as malicious only when supported by evidence such as unauthorized access, an invalid command source, a privilege mismatch, or an unexpected equipment response. This context-aware procedure reduced the possibility of treating legitimate operational actions as cyberattacks. Normal records represented authorized communication and expected variations in grid operation [19]. They included routine demand fluctuations, scheduled switching, approved maintenance, renewable-generation changes, and temporary network variations that remained within accepted operational conditions. Clearly defining the normal class was important because legitimate grid behaviour may sometimes resemble an attack. Seasonal consumption variation, for example, could appear similar to energy manipulation if historical and contextual information were ignored. The multi-class labelling workflow is presented in Figure 4.

Integrated smart-grid events were first examined for data integrity, communication, access-control, identity, and malware-related evidence [20]. Relevant events were then correlated across operational and cybersecurity sources before receiving a final class label. Records containing insufficient or contradictory evidence were placed in a review group and excluded from supervised model training when their class could not be reliably determined.

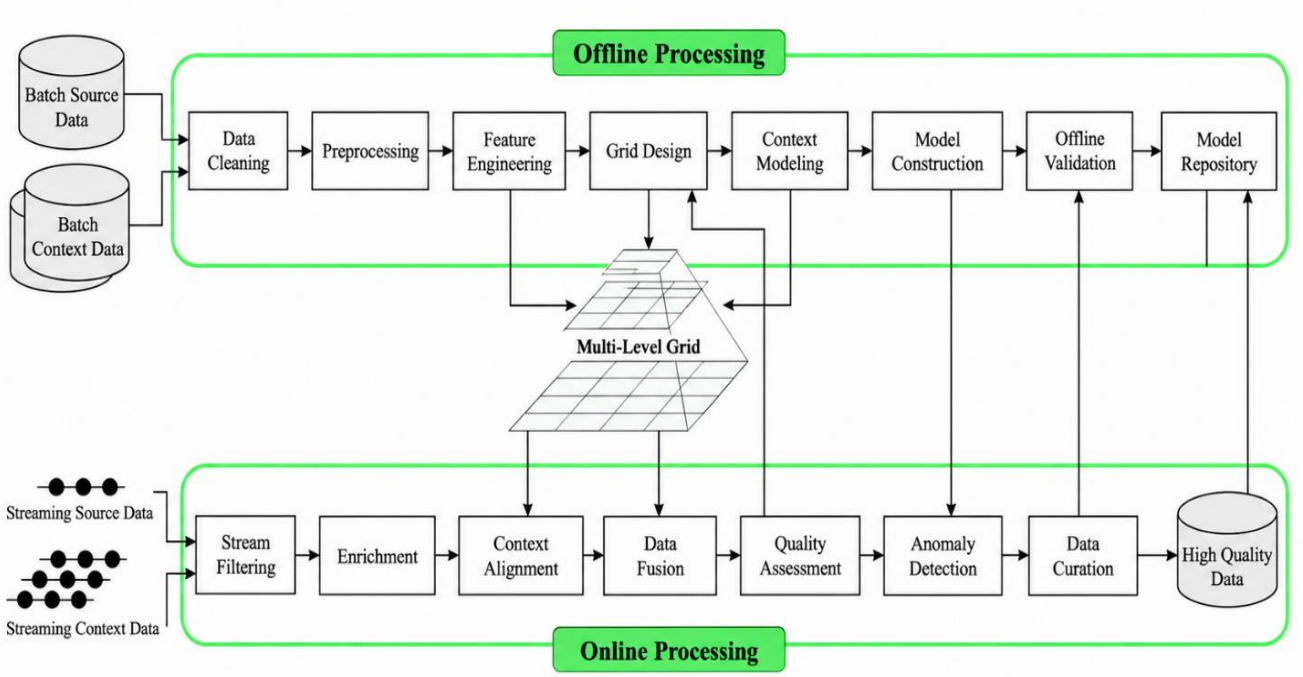

Figure 4: Integrated offline–online framework for multi-level contextual data processing and quality enhancement.

Temporal consistency was also considered during labelling. Related observations occurring within the same attack window were grouped to prevent individual stages of a coordinated event from being assigned conflicting labels. When multiple indicators were present, the final label was based on the attack's dominant mechanism and verified operational effect. This approach supported the identification of multistage attacks while maintaining a clear single-label structure for supervised classification. The completed threat library provided standardized definitions and evidence requirements for every class. It also reduced subjective labelling variation and supported consistent comparison across baseline and proposed models [21]. By combining attack signatures with operational context and cross-source event correlation, the resulting labels provided a reliable foundation for training the CNN–BiLSTM–Attention model to distinguish normal behaviour from six important smart-grid cyberattack categories.

4.3- Lightweight CNN–BiLSTM–Attention Architecture:

A lightweight CNN–BiLSTM–Attention architecture was developed to detect normal activity and six cyberattack categories within the smart-grid IoT–cloud ecosystem. The architecture was designed to capture three complementary forms of information: local relationships among predictive variables, temporal dependencies across consecutive events, and the contextual importance of individual observations. Computational complexity was controlled so that the trained model could perform real-time inference on

resource-constrained edge gateways. The model received synchronized event sequences containing the 48 selected predictive variables. These variables represented electrical measurements, device conditions, network activity, authentication behaviour, and energy-consumption patterns. Each sequence was first processed by lightweight one-dimensional convolutional layers. These layers extracted local patterns such as voltage–current inconsistencies, sudden packet-rate changes, unusual login combinations, and abnormal consumption variations [22]. Batch normalization stabilized the extracted representations, while nonlinear activation supported the learning of complex relationships. A pooling layer reduced feature-map dimensions and lowered the computational cost of subsequent processing. The CNN output was transferred to bidirectional long short-term memory layers. The forward component analysed events in chronological order, whereas the backward component examined the same sequence in the reverse direction. Combining both representations allowed the model to interpret suspicious activity according to preceding and subsequent events. This capability was particularly useful for detecting gradual false-data manipulation, repeated access attempts, coordinated denial-of-service behaviour, and multistage malware intrusion. An attention layer then evaluated the importance of the BiLSTM representations [23]. Greater attention was assigned to the events and time intervals that provided the strongest evidence of malicious activity. The resulting context-rich representation was passed to fully connected layers for final classification. The Softmax output layer produced probabilities for seven classes: normal activity, false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. The roles of the principal architectural components are summarized in Table 5.

Table 5: Components of the lightweight CNN–BiLSTM–Attention architecture

Architectural component	Primary function	Lightweight design consideration
Input sequence layer	Receives synchronized event windows	Fixed-length sequences and 48 selected variables
One-dimensional CNN	Extracts local feature relationships	Limited filters and compact kernel sizes
Batch normalization	Stabilizes intermediate representations	Low additional computational cost
Pooling layer	Reduces representation dimensions	Decreases memory use and processing time
BiLSTM layer	Learns bidirectional temporal dependencies	Controlled number of hidden units
Attention layer	Prioritizes important events and time steps	Compact attention mechanism
Fully connected layer	Integrates learned representations	Reduced layer dimensions
Softmax output layer	Generates seven-class probabilities	Single final classification layer

The complete processing sequence is illustrated in Figure 5. Synchronized smart-grid event windows enter the CNN component for local feature extraction. The resulting feature maps are compressed and forwarded to the BiLSTM layer for temporal analysis. Attention-based prioritization highlights the most relevant representations, after which the classification layers produce the predicted threat class and confidence score.

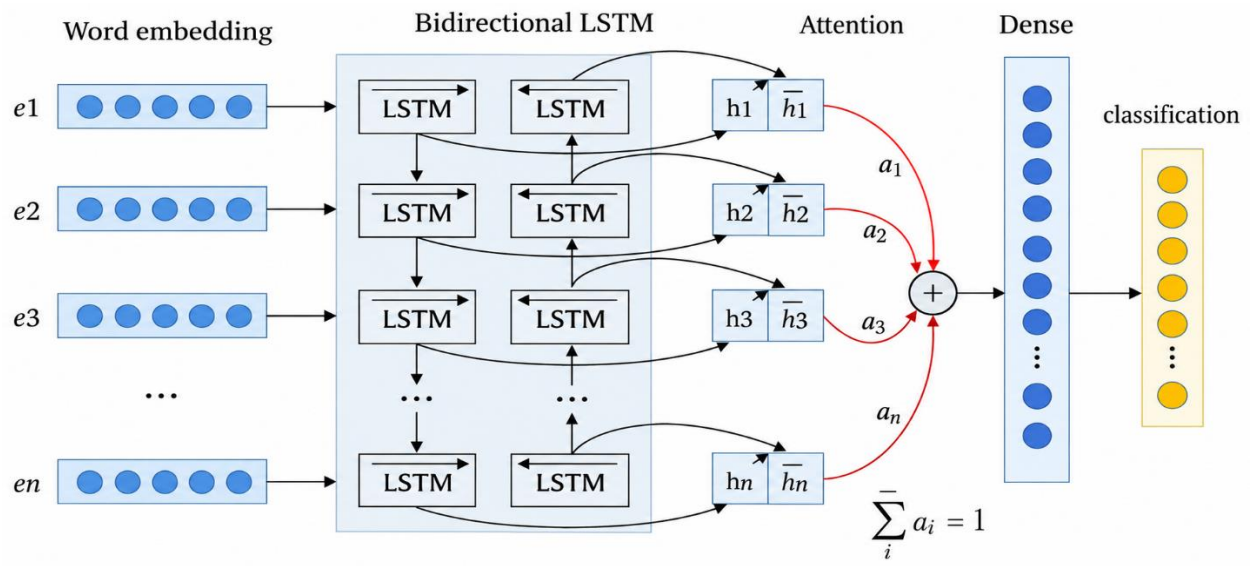

Figure 5: Proposed lightweight CNN–BiLSTM–Attention architecture for smart-grid cyberattack detection.

Several design controls were applied to improve suitability for edge deployment. The number of convolutional filters and BiLSTM units was restricted to reduce model size. Pooling minimized the dimensions of intermediate representations, while dropout reduced overfitting without increasing inference complexity. Input sequences were limited to an operationally useful length so that temporal information could be preserved without generating excessive computational delay. Model configurations were evaluated according to validation performance, parameter count, memory demand, and inference time. The hybrid architecture offered advantages over individual deep-learning components [24]. A conventional CNN could detect local anomalies but might overlook the sequence in which events occurred. An LSTM-based model could analyse temporal behaviour but might not extract detailed interactions among heterogeneous variables as effectively. The attention mechanism further improved the integrated model by reducing the influence of routine observations and prioritizing attack-related evidence. Therefore, the proposed architecture supported accurate detection of both immediate and gradually developing cyberattacks. After training, the optimized model was deployed at smart-grid edge gateways for local threat detection. Each prediction included an attack class and confidence level, which were forwarded to the adaptive response component [25]. High-confidence and high-risk detections could initiate immediate protective actions, whereas uncertain events were escalated for additional validation. This lightweight architecture consequently provided the analytical foundation for accurate, low-latency, and context-aware security across distributed smart-grid IoT–cloud environments.

4.4- Real-Time Edge Inference and Cloud Coordination:

The proposed framework distributed cybersecurity functions between local edge gateways and the cloud according to their latency, privacy, and computational requirements. Edge nodes were positioned close to smart meters, PMUs, substation sensors, intelligent electronic devices, and communication gateways. These nodes performed time-sensitive operations, including data preprocessing, sequence generation, attack detection, confidence estimation, and initial threat response. The cloud layer was responsible for computationally intensive and non-urgent tasks such as federated model aggregation, long-term trend analysis, model version management, historical data storage, and security-policy

coordination. Incoming records were first processed at the edge using the same cleaning, feature-transformation, and normalization procedures applied during model training. Temporally related observations were arranged into fixed-length event sequences containing the 48 selected variables [26]. The optimized CNN–BiLSTM–Attention model then analysed each sequence and generated a predicted class and confidence score. Predictions covered normal behaviour and six attack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. When normal activity or a low-risk anomaly was detected, the edge node retained the detailed records locally and transmitted only compact summaries to the cloud. High-confidence attacks were immediately forwarded to the adaptive threat-response module, allowing protective actions to be initiated without waiting for cloud authorization when predefined emergency conditions were satisfied. Uncertain or operationally sensitive cases were escalated to the cloud or a human security analyst for additional validation. Cloud coordination supported the exchange of global model parameters rather than raw consumer and grid-operational data. During federated training, authorized edge nodes received the current model, performed local updates, and returned protected parameters for aggregation. Differential privacy and secure communication were used to reduce information exposure during this exchange [27]. After aggregation, the updated global model was validated before being redistributed to participating edge nodes. Model version controls ensured that an unstable or corrupted update could be rejected or replaced by an earlier trusted version. The real-time inference and coordination workflow is illustrated in Figure 6. Smart-grid data were analysed locally at the edge, after which only security alerts, compressed summaries, and protected model updates were communicated to the cloud. Updated models and coordinated security policies were returned to the edge nodes for continued local operation.

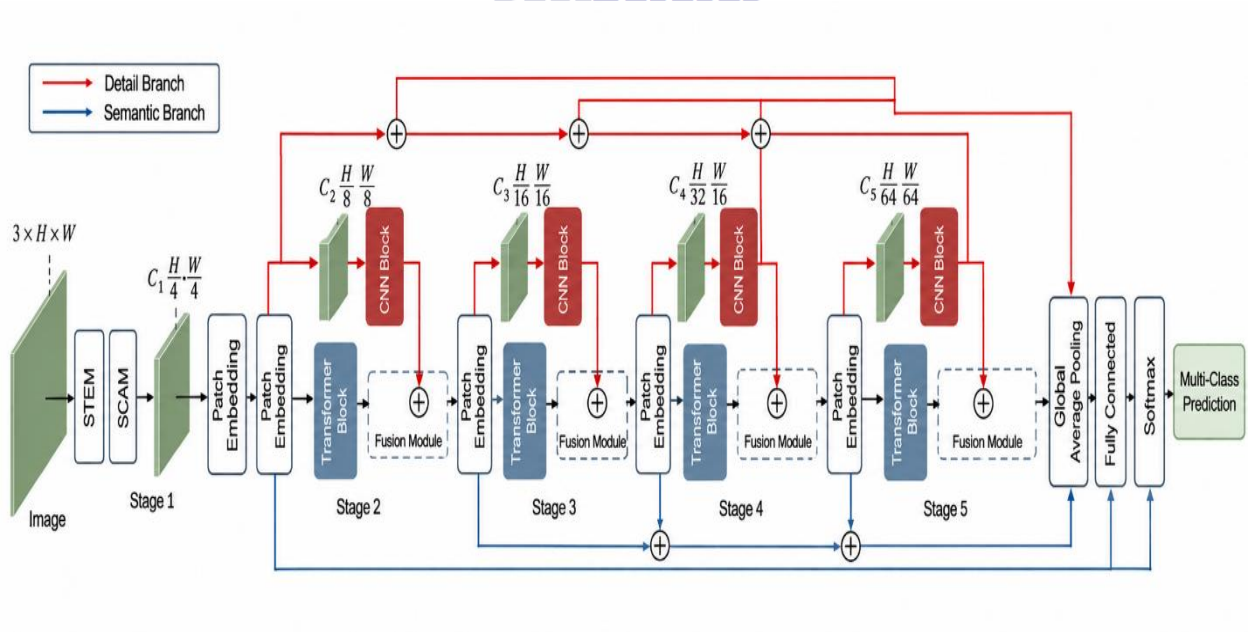

Figure 6: Real-time edge inference and cloud-coordination workflow.

The architecture was also designed to maintain essential cybersecurity functions during intermittent cloud connectivity. If communication with the cloud was unavailable, the edge node continued to preprocess incoming data, execute the local detection model, and apply approved response policies. Security summaries and model updates were temporarily queued and

transmitted after connectivity was restored. This local autonomy reduced dependence on continuous cloud availability and prevented temporary communication failures from disabling the threat-detection system [28]. Operational performance was evaluated by measuring end-to-end detection latency and the volume of cloud-bound data. Detection latency covered the interval between the arrival of a complete event sequence and generation of the corresponding threat prediction. The edge-based configuration was compared with a cloud-only arrangement in which records had to be transmitted to a remote platform before analysis. The evaluation showed that average detection latency decreased from 186 ms under cloud-only processing to 42 ms with edge inference, representing a 77.4% reduction. Communication efficiency was assessed by comparing the quantity of raw data transmitted in the cloud-only configuration with the volume of summaries, alerts, and model updates exchanged by the proposed architecture [29]. Local processing reduced cloud-bound data transmission by 68.2%. These improvements demonstrate that coordinated edge and cloud processing can provide rapid attack detection while limiting bandwidth demand and reducing the exposure of sensitive energy information.

4.5- Risk-Aware Adaptive Threat Response:

A risk-aware adaptive threat-response mechanism was incorporated to translate cyberattack predictions into timely and proportionate security actions. Instead of treating every detected anomaly equally, the mechanism evaluated each event according to its predicted attack class, model confidence, attack severity, affected asset, recurrence pattern, and potential operational impact. This contextual assessment reduced unnecessary interventions while enabling rapid containment of threats capable of disrupting critical smart-grid functions. The response mechanism received the predicted class and confidence score generated by the CNN–BiLSTM–Attention model. This information was combined with operational factors such as device criticality, anomaly intensity, communication behaviour, previous security events, and the number of interconnected assets potentially exposed to the attack. A suspicious login associated with a consumer-level smart meter, for example, did not receive the same priority as an unauthorized command directed toward a substation breaker or protection relay [30]. Detected events were classified into four response levels: low, medium, high, and critical. Low-risk events were recorded and monitored for repeated abnormal behaviour. Medium-risk incidents initiated enhanced monitoring, additional authentication, or temporary communication restrictions. High-risk events generated priority alerts and could restrict malicious traffic or suspend suspicious sessions. Critical events activated immediate containment measures, including device isolation, command blocking, credential revocation, trusted-channel switching, and operator notification. The risk levels and associated responses are summarized in Table 6.

Table 6: Risk levels and adaptive response actions for detected smart-grid threats

Risk level	General decision criteria	Human involvement
Low	Low-confidence anomaly with limited operational impact	Periodic analyst review
Medium	Repeated anomaly or moderate-confidence attack affecting a non-critical asset	Analyst notification
High	High-confidence attack or suspicious activity affecting an important device	Rapid operator validation
Critical	Confirmed or high-confidence attack threatening essential grid operations	Immediate operator notification and oversight

Response selection was attack-specific. False-data injection incidents triggered measurement validation against neighbouring sensors and trusted state estimates. Denial-of-service attacks initiated traffic filtering, rate limiting, and rerouting through available communication paths. Energy-theft events were forwarded for meter verification and billing review. Command-injection incidents caused suspicious instructions to be blocked and the associated access session to be suspended. Device-spoofing detections initiated identity revalidation and temporary communication isolation, while malware-based intrusions activated device quarantine, process restriction, and forensic logging. The figure 7 illustrate the Self-attention-based BiLSTM architecture for polarity and sentiment classification. A detected event first underwent contextual risk assessment [31]. The system then selected a response level and executed the corresponding containment measures. Outcomes were monitored to determine whether the threat had been controlled. If suspicious activity continued, the event was escalated to a higher response level and forwarded for human review.

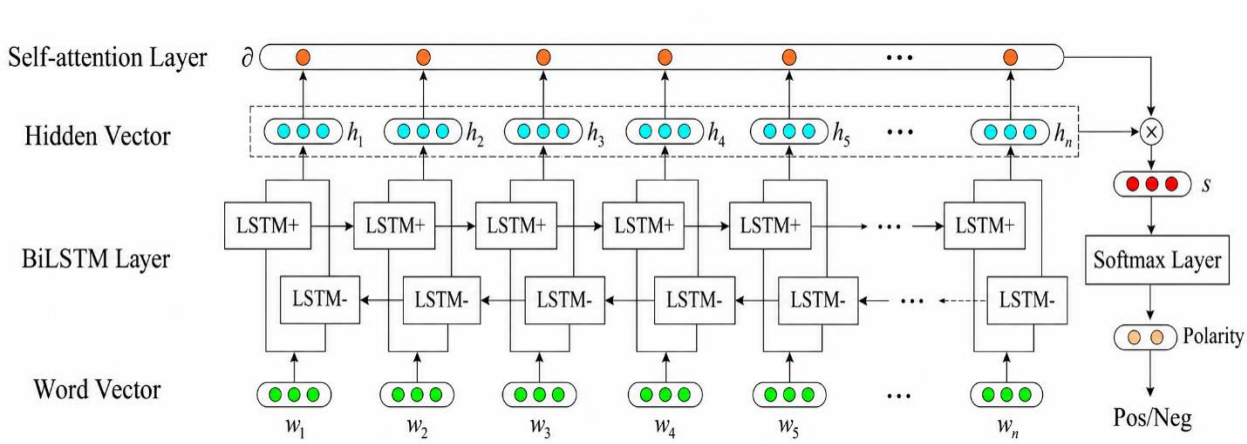

Figure 7: Self-attention-based BiLSTM architecture for polarity and sentiment classification.

Human-in-the-loop oversight was retained for actions capable of affecting service continuity or safety-critical equipment. Although the edge node could automatically block clearly malicious traffic or revoke an unauthorized session, actions such as disconnecting a protection device or modifying a critical controller required operator confirmation unless predefined emergency conditions were met. This arrangement balanced rapid automation with the operational reliability required in critical power infrastructure. The mechanism also maintained response logs containing the detected attack type, affected asset, confidence level, selected action, execution time, and final outcome [32]. These records supported forensic investigation, policy refinement, and future model updates. Repeated incidents and unsuccessful responses were used to modify alert priorities and strengthen subsequent containment decisions. The effectiveness of the mechanism was evaluated by comparing successful attack propagation before and after adaptive response activation. Attack propagation referred to the ability of an intrusion to spread from its initial entry point to additional devices, services, or operational zones. Experimental results showed that the proposed response process reduced successful attack propagation by 71.6%. This finding demonstrates that combining intelligent detection with contextual risk assessment and proportionate containment can substantially strengthen the resilience of smart-grid IoT–cloud ecosystems.

5- Results and Discussion:

The proposed CNN–BiLSTM–Attention model demonstrated strong performance in

distinguishing normal smart-grid activity from six cyberattack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. On the independent test dataset, the model achieved 98.1% accuracy, 97.8% precision, 97.6% recall, a 97.7% F1-score, and a ROC–AUC of 0.993. These findings indicate that the model maintained a favourable balance between identifying malicious events and limiting false alarms. The high recall is particularly important for critical infrastructure because undetected attacks may manipulate grid measurements, compromise control devices, or interrupt communication. At the same time, the high precision reduces unnecessary investigations and prevents excessive security responses to legitimate operational changes. Table 7 present the Overall cyberattack-detection performance of the proposed model.

Table 7: Overall cyberattack-detection performance of the proposed model

Performance metric	Result
Accuracy	98.1%
Precision	97.8%
Recall	97.6%
F1-score	97.7%
ROC–AUC	0.993

The ROC–AUC value of 0.993 confirms that the framework provided excellent discrimination between normal and malicious events across different classification thresholds. This flexibility is valuable in smart-grid environments because security thresholds can be adjusted according to asset criticality. A lower threshold may be selected for protection relays, substations, and control gateways to minimize missed attacks, whereas a more restrictive threshold may be applied to less critical devices to limit false alerts [33]. The proposed architecture outperformed all baseline models, as presented in Table 8. Logistic regression achieved an accuracy of 84.7%, indicating that linear relationships were insufficient to represent the complex interactions among electrical, network, authentication, and behavioural variables. The support vector machine improved accuracy to 89.6% because of its nonlinear classification capability. Random forest achieved 93.8%, demonstrating the benefit of ensemble-based feature learning, whereas the conventional CNN reached 95.4% by extracting local patterns from synchronized smart-grid sequences.

Table 8: Accuracy comparison of the proposed framework and baseline models

Model	Accuracy	Difference from proposed model
Logistic regression	84.7%	−13.4 percentage points
Support vector machine	89.6%	−8.5 percentage points
Random forest	93.8%	−4.3 percentage points
Conventional CNN	95.4%	−2.7 percentage points
Proposed CNN–BiLSTM–Attention	**98.1%**	–

The proposed model improved accuracy by 13.4 percentage points over logistic regression, 8.5 points over the support vector machine, 4.3 points over random forest, and 2.7 points over the conventional CNN. The improved performance can be attributed to the complementary functions of its architectural components. CNN layers captured local relationships among the 48 predictive variables, including measurement inconsistencies, packet-rate changes, identity conflicts, and consumption deviations [34]. BiLSTM layers examined the progression of events

in forward and backward temporal directions, while the attention mechanism emphasized the observations that provided the strongest evidence of malicious behaviour. The multi-source structure of the dataset further strengthened attack recognition. False-data injection was identified through inconsistencies among sensor readings, PMU measurements, and expected electrical conditions. Denial-of-service attacks generated distinctive network patterns involving abnormal packet frequency, repeated requests, connection failures, and communication delays. The figure 8 compares training accuracy over 200 communication rounds for 8-bit, 16-bit, and 24-bit quantization against a non-quantized model. All configurations rapidly converge toward approximately 96–98% accuracy, although the 8-bit model exhibits greater fluctuations. The 16-bit and 24-bit approaches remain close to the non-quantized baseline, indicating that moderate quantization can reduce computational and communication costs while largely preserving predictive performance.

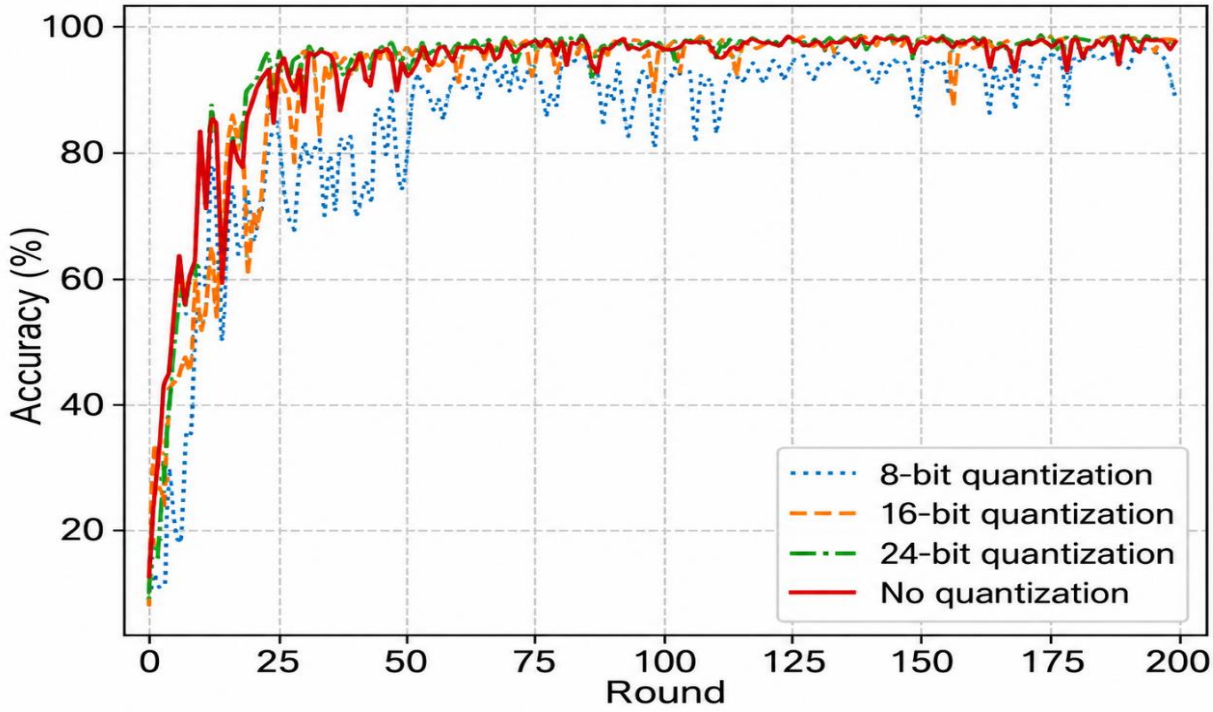

Figure 8: Accuracy convergence across communication rounds under different model-quantization levels.

Energy-theft detection relied on irregular relationships among meter states, historical consumption, and observed load profiles. Command-injection attacks were recognized through unauthorized instructions, access-privilege conflicts, and unexpected equipment responses. Device-spoofing detection benefited from the joint analysis of device identity, access origin, authentication outcomes, and communication behaviour. Malware-based intrusions required broader temporal analysis because their indicators could develop gradually through abnormal processes, persistent connections, privilege escalation, and lateral movement [35]. These findings demonstrate that integrating multiple data sources provides more reliable evidence than examining electrical measurements or network traffic independently. The difference between the conventional CNN and the proposed model highlights the importance of temporal learning and contextual prioritization. Although the CNN effectively

identified local anomalies, it could not fully represent the order and duration of related security events. The BiLSTM component allowed the model to interpret each event according to its preceding and subsequent context. For example, an individual successful login might appear normal, but it becomes suspicious when preceded by repeated authentication failures and followed by unauthorized control activity. The attention mechanism further improved classification by reducing the influence of routine measurements and prioritizing unusual commands, measurement deviations, traffic bursts, and identity inconsistencies.

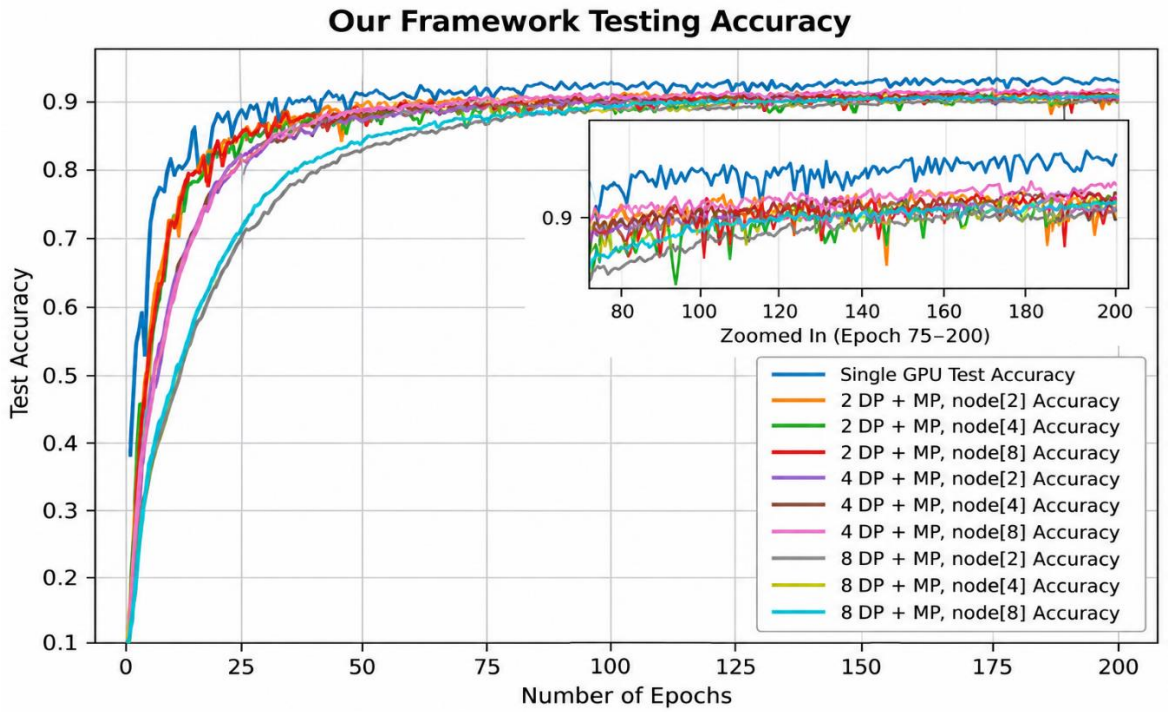

Figure 9: Comparative test-accuracy convergence of the proposed framework under different data- and model-parallel configurations.

The figure 9 presents test accuracy over 200 training epochs for a single-GPU baseline and multiple distributed configurations combining data parallelism (DP) and model parallelism (MP) across 2, 4, and 8 nodes. All configurations improve rapidly during the initial epochs and stabilize at approximately 0.89–0.93 accuracy. The single-GPU model achieves the highest final accuracy, while the distributed configurations demonstrate comparable and stable convergence. The inset highlights minor accuracy fluctuations between epochs 75 and 200, confirming the framework's consistent performance across different computational settings. The framework also produced substantial improvements in real-time operational performance. In the cloud-only configuration, average end-to-end threat-detection latency was 186 ms. Edge deployment reduced this value to 42 ms, representing a 77.4% reduction. This improvement occurred because data were analysed close to their originating devices without waiting for complete transmission to a remote cloud platform. The reduction is particularly important for denial-of-service, command-injection, and false-data injection attacks, where even a short delay may allow malicious activities to influence additional devices or operational processes. Local processing also reduced cloud-bound data transmission by 68.2%. Edge nodes transmitted security alerts, compact summaries, and protected model updates instead of continuously uploading raw measurements, authentication logs, and energy profiles [36]. Only 31.8% of the data volume associated with the

cloud-only configuration therefore required transmission. This reduction decreased bandwidth demand, limited the exposure of sensitive information, and allowed edge nodes to maintain essential detection functions during temporary cloud-connection failures. Table 9 shows the Operational and privacy-related performance of the framework.

Table 9: Operational and privacy-related performance of the framework

Evaluation indicator	Baseline condition	Proposed condition	Observed improvement
Average detection latency	186 ms	42 ms	77.4% reduction
Cloud-bound data transmission	100%	31.8%	68.2% reduction
Successful attack propagation	100%	28.4%	71.6% reduction
Detection accuracy with privacy protection	98.1% non-private accuracy	96.9% private accuracy	1.2-percentage-point reduction

Differentially private federated learning preserved decentralized consumer and grid-operational information with only a limited reduction in model performance. The non-private model achieved 98.1% accuracy, whereas the differentially private configuration maintained approximately 96.9%, corresponding to a reduction of 1.2 percentage points. This result indicates that privacy protection can be introduced without making the detection model operationally ineffective. Federated learning allowed distributed edge nodes to contribute to model development without transmitting raw local datasets, while differential privacy reduced the possibility of inferring sensitive information from shared updates. The figure 10 compares the proposed framework with its non-PPO variant, FedProx-Async, DBAFL, FedProx, FedAsync, FedAsyncc, and FedAvg under (a) 0%, (b) 5%, (c) 10%, and (d) 30% malicious-node conditions. All methods perform similarly in attack-free settings, whereas accuracy and convergence stability decline as the proportion of malicious nodes increases. The proposed framework consistently achieves the highest or near-highest accuracy, particularly under 10% and 30% malicious-node scenarios, demonstrating stronger robustness and more stable convergence than the baseline methods.

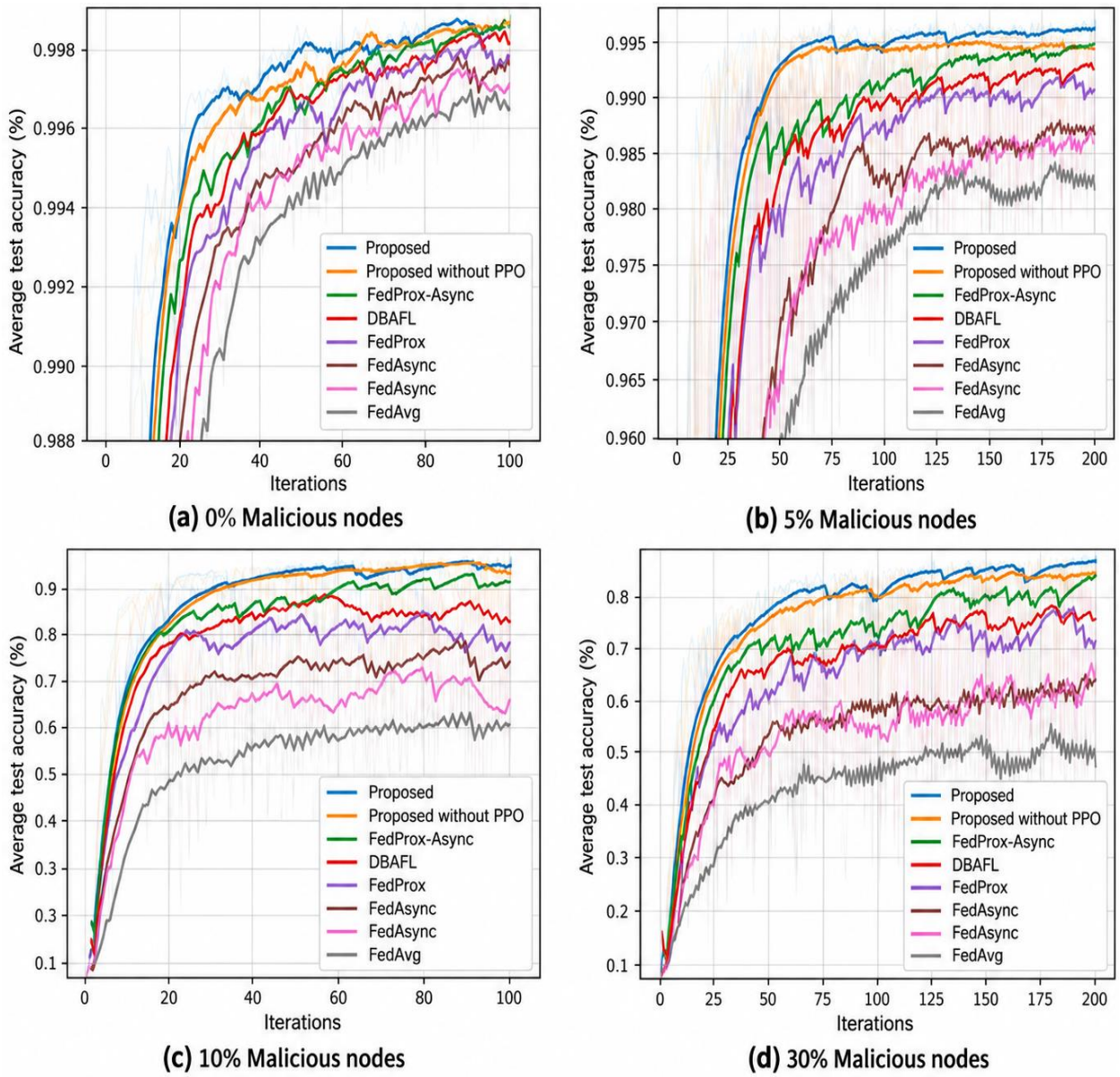

Figure 10: Comparative test-accuracy convergence of the proposed framework and federated-learning baselines under varying proportions of malicious nodes.

The adaptive threat-response mechanism reduced successful attack propagation by 71.6%, leaving the propagation level at 28.4% of the uncontrolled baseline. Rather than applying the same response to every event, the mechanism considered the predicted attack category, confidence level, asset criticality, anomaly intensity, and recurrence pattern. Denial-of-service events initiated traffic filtering and rate limiting, while device-spoofing incidents triggered identity revalidation and communication restriction. Suspicious commands were blocked, compromised credentials were revoked, and malware-affected devices were isolated for investigation. Risk-based response helped prevent low-confidence anomalies from initiating unnecessarily disruptive actions [37]. Low-risk events were logged and monitored, medium-risk events received additional verification, and high or critical threats activated immediate containment measures. Human

oversight was retained for decisions involving protection devices, substation equipment, and other safety-critical infrastructure. This balance between automated response and human control is necessary because an inappropriate containment action could disrupt legitimate electricity delivery. Although the results demonstrate considerable advantages, several limitations remain. The evaluation addressed six major attack classes and may not represent every emerging or zero-day threat. Differences in grid topology, communication protocols, device hardware, and consumer behaviour may also influence performance in other deployment environments. The figure 11 compares the z-axis acceleration responses produced by the FP and MP methods during a two-second interval. Both profiles exhibit similar initial deceleration and subsequent positive acceleration, followed by a sustained negative phase near −10 m/s². At approximately 1.45 s, the FP signal shows a sharp transient peak of nearly 41 m/s², whereas the MP profile produces a smoother peak of about 17 m/s². Both responses then gradually decrease and approach zero, highlighting differences in transient sensitivity and response smoothness.

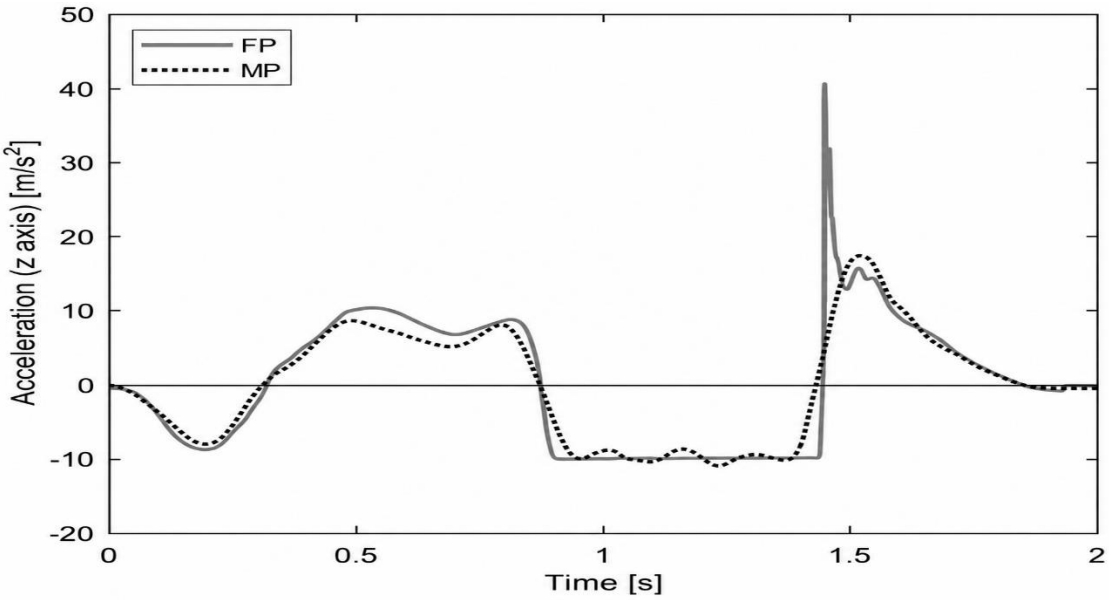

Figure 11: Comparative *z*-axis acceleration profiles of the FP and MP methods over time.

Federated learning may face non-uniform local data distributions, unreliable client participation, or malicious model updates. Furthermore, differential privacy introduces an unavoidable trade-off between information protection and classification accuracy [38]. Overall, the findings confirm that the proposed framework provides accurate, low-latency, privacy-aware, and context-sensitive cybersecurity for smart-grid IoT–cloud ecosystems. Its 98.1% accuracy and 0.993 ROC–AUC exceeded the performance of the selected baseline models. Edge inference reduced detection latency by 77.4% and cloud-bound transmission by 68.2%, while the adaptive response mechanism reduced successful attack propagation by 71.6%. Combined with the limited 1.2-percentage-point accuracy cost of privacy protection, these results demonstrate the practical value of integrating lightweight deep learning, edge computing, federated learning, differential privacy, and adaptive response within a unified smart-grid security framework.

6- Future Work:

Future research should evaluate the proposed Edge AI framework across large-scale, geographically distributed smart-grid environments containing diverse devices, communication protocols, network topologies, and operational conditions. Hardware-in-the-loop

experiments and real-world pilot deployments would provide further evidence regarding inference speed, memory consumption, energy requirements, communication reliability, and long-term operational stability [39]. The framework should also be tested under unstable connectivity, edge-node failure, rapid load variation, and high penetration of distributed energy resources and electric vehicles. The existing threat library can be expanded beyond the six evaluated attack categories to include zero-day attacks, ransomware, advanced persistent threats, coordinated multi-vector intrusions, adversarial manipulation, model-poisoning attacks, and insider threats [40]. Future models could incorporate continual and online learning to identify emerging attack behaviour without requiring complete retraining. Concept-drift detection would also allow the system to distinguish new cyberattack patterns from legitimate changes in consumption, renewable generation, and grid configuration. Further investigation is required to improve federated learning under non-uniform data distributions and unreliable edge-node participation. Secure aggregation, robust client selection, poisoning-resistant model aggregation, and blockchain-supported update verification could strengthen trust among participating grid operators. Privacy-budget optimization and adaptive noise mechanisms may further protect sensitive consumer information while reducing the loss in detection accuracy caused by differential privacy.

Future studies should also explore model compression, pruning, quantization, knowledge distillation, and specialized edge accelerators to reduce processing and memory requirements. Explainable AI techniques could be integrated to provide clearer evidence for each alert, including influential variables, suspicious time intervals, and relationships among linked events. Digital-twin environments could be used to simulate alternative response strategies before applying them to physical grid assets [41]. Finally, the adaptive threat-response component should be validated against a wider range of operational scenarios and regulatory requirements. Reinforcement learning may enable the system to select containment actions dynamically according to attack severity, asset criticality, and expected operational consequences. However, human oversight, fail-safe controls, and formal policy constraints should remain central to all safety-critical decisions. These extensions could support the development of a more autonomous, resilient, explainable, and privacy-aware cybersecurity architecture for future smart grids.

Conclusion:

This study developed a real-time Edge AI framework for securing smart-grid IoT–cloud ecosystems through intelligent cyberattack detection, privacy-preserving energy analytics, and adaptive threat response. The framework integrated heterogeneous information from smart meters, phasor measurement units, substation sensors, network traffic, authentication logs, and energy-consumption profiles. A total of 285,000 operational and cybersecurity records were processed, from which 48 predictive variables were retained. The resulting dataset represented normal grid behaviour and six major attack categories: false-data injection, denial-of-service, energy theft, command injection, device spoofing, and malware-based intrusion. The proposed lightweight CNN–BiLSTM–Attention architecture combined local feature extraction, bidirectional temporal analysis, and context-sensitive threat prioritization. It achieved 98.1% accuracy, 97.8% precision, 97.6% recall, a 97.7% F1-score, and a ROC–AUC of 0.993. Its accuracy exceeded those of logistic regression, support vector machine, random forest, and conventional CNN models, which achieved 84.7%, 89.6%, 93.8%, and 95.4%, respectively. These findings demonstrate that integrating spatial, temporal, and attention-based learning improves the identification of complex and evolving smart-grid attacks. Deploying the optimized model at the network edge reduced average detection latency from 186 ms to 42 ms, representing a 77.4% improvement, while decreasing cloud-bound data transmission by 68.2%. Federated learning enabled collaborative model development without centralizing raw consumer or grid-operational data, and differential privacy limited sensitive

information exposure with only a 1.2-percentage-point reduction in detection accuracy. Furthermore, the risk-aware response mechanism isolated compromised devices, prioritized security alerts, and reduced successful attack propagation by 71.6%. Overall, the findings confirm that combining lightweight deep learning, edge computing, federated learning, differential privacy, and adaptive response can provide accurate, rapid, scalable, and privacy-aware protection for next-generation smart grids. Although additional validation is required under large-scale real-world conditions and emerging attack scenarios, the proposed framework offers a strong foundation for resilient cybersecurity across increasingly distributed and interconnected energy infrastructures.

REFERENCES:

Sanjalawe, Y., Fraihat, S., Makhadmeh, S. N., & Alzubi, E. (2025). AI-powered smart grids in the 6G era: A comprehensive survey on security and intelligent energy systems. *IEEE Open Journal of the Communications Society*.

Alatawi, M. N. (2025). EdgeGuard-IoT: 6G-enabled edge intelligence for secure federated learning and adaptive anomaly detection in Industry 5.0. *Computers, Materials & Continua*, *85*(1), 695-727.

Ahmad, B., Ali, S., Muneer, M., Fatima, A., & Abbas, N. (2025). Wind energy integration into the SAARC region: A comprehensive review of optimal wind sites for a sustainable super smart grid. Wind Energy, 3(2).

Molokomme, D. N., Onumanyi, A. J., & Abu-Mahfouz, A. M. (2022). Edge intelligence in Smart Grids: A survey on architectures, offloading models, cyber security measures, and challenges. *Journal of Sensor and Actuator Networks*, *11*(3), 47.

Zhukabayeva, T., Zholshiyeva, L., Karabayev, N., Khan, S., & Alnazzawi, N. (2025). Cybersecurity solutions for industrial internet of things–edge computing integration: Challenges, threats, and future directions. *Sensors*, *25*(1), 213.

Akinsanya, M. O., Bello, A. B., & Adeusi, O. C. (2023). A comprehensive review of edge computing approaches for secure and efficient data processing in IoT networks. *Communication in Physical Sciences*, *9*(4), 870-877.

Ahmad, B., Majeed, M. K., Soomro, A. A., & Jillani, S. A. (2026, January). Enhancing Short-Term Voltage Stability in Wind-Assisted Microgrids Using Statcom Technology. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Rahmati, M., & Rahmati, N. (2025). Adaptive federated edge intelligence for real-time cyberthreat detection in resource-constrained iot environments: a lightweight deep learning approach. *Journal of Computer Virology and Hacking Techniques*, *21*(1), 35.

Gollavilli, V. S. B. H. (2025). Ai-optimized cloud-edge collaborative systems for data privacy: Statistical detection, pca, gwo integration, and fog computing. *International Journal of Automation and Smart Technology*, *15*(1).

Mrabet, M., & Sliti, M. (2025). Towards secure, trustworthy and sustainable edge computing for smart cities: Innovative strategies and future prospects. *IEEe Access*.

Sousa, A., Correia, L., & Reis, M. J. (2025, May). Edge-based AI for real-time threat detection in 5G-IoT networks: A comparative and architectural review. In *2025 5th Intelligent Cybersecurity Conference (ICSC)* (pp. 359-363). IEEE.

Ali, S., Alam, F., Hafeez, S., & Khan, A. (2025). Artificial Intelligence and Big Data–Enabled Architectures for Privacy-Preserving and Cybersecure Demand Response in Smart Grids.

Narayan, R. K. (2025). Cybersecurity Challenges in Cloud-Edge Computing Convergence: A Systematic Analysis and Adaptive Defense Framework. *International Journal of Cyberspace Security*, *1*(1), 38-50.

Takruri, M., Rabih, M., Dbeiss, L. M., Shall, H., Badawi, S. A., Al Atem, M., & Arnaout, M. (2026). AI–IoT-Enabled Smart Energy Ecosystems: Architectures, Security, and Sustainability. *Eng*, *7*(7), 354.

Albshaier, L., Almarri, S., & Albuali, A. (2025). Federated learning for cloud and edge security: A systematic review of challenges and AI opportunities. *Electronics*, *14*(5), 1019.

Khan, P. A. (2026). AI-Driven Edge Computing Architecture for Real-Time Data Analysis in Next-Generation Smart Applications. *Journal of Smart Computing and Data Intelligence*, *2*(1), 33-41.

Ahmad, B., Jillani, S. A., Rafique, M., & Soomro, A. A. (2026, January). Design and Monitoring of a Tree-Inspired Vertical Axis Wind Turbine for Efficient Urban Wind Utilization. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Alao, G. U., Onyema, K., & Adeoba, M. I. (2026). Privacy-Preserving Intelligence for Electric Vehicle Charging Networks: Federated Learning, Edge Analytics, and Grid Integration. *Journal of Trending Research*, *1*(1), 17-25.

Yıldırım, F., Yalman, Y., Bayındır, K. Ç., & Terciyanlı, E. (2025). Comprehensive review of edge computing for power systems: State of the art, architecture, and applications. *Applied Sciences*, *15*(8), 4592.

Eren, H., Karaduman, Ö., & Gençoğlu, M. T. (2025). Security and privacy in the Internet of Everything (IoE): A review on blockchain, edge computing, AI, and quantum-resilient solutions. *Applied Sciences*, *15*(15), 8704.

Arif, A., Ali, M. Z. H., Haider, S. Z. Q., & Niaz, Q. (2025). AI-Driven Edge Computing for IoT: Revolutionizing Phishing Detection and Mitigation. *J. Comput. Biomed. Inf*, *9*(1), 1-15.

Salim, I., Mughal, U. A., Naveed, O., & Hafeez, S. (2025). Toward Secure and Privacy-Preserving Smart Grids: Communication Architectures, Cybersecurity Innovations, and Policy Frameworks. *Spectrum of Engineering Sciences*, 1188-1232.

Akbar, R., & Zafer, A. (2024). Next-gen information security: AI-driven solutions for real-time cyber threat detection in cloud and network environments. *J. Cybersecur. Res*, *12*, 123-145.

Wang, C., Yuan, Z., Zhou, P., Xu, Z., Li, R., & Wu, D. O. (2023). The security and privacy of mobile-edge computing: An artificial intelligence perspective. *IEEE Internet of Things Journal*, *10*(24), 22008-22032.

Lakhal, H., Zegrari, M., & Bahnasse, A. (2025). Next-generation smart grid cybersecurity: A systematic review of OT cyber threats, AI-driven defense, cyber deception techniques, and emerging security strategies. *IEEE Access*.

Ali, S. (2024). Securing edge computing with AI: Intelligent threat detection for decentralized systems. *Artificial Intelligence in Edge Security*; *ResearchGate: Berlin, Germany*.

Angle, M. (2024). Revolutionizing Data Forensics: Leveraging Edge Computing for Enhanced Security in IoT Ecosystems.

Hameed, K., Maqsood, F., Patwary, M. A. K., Zhenfie, W., & Sarwar, R. A Scalable Privacy-Preserving Artificial Intelligence Framework for Real-Time Threat Detection in Internet of Things Systems. *Available at SSRN 5228905*.

Ari, I., Balkan, K., Pirbhulal, S., & Abie, H. (2024, December). Ensuring security continuum from edge to cloud: Adaptive security for iot-based critical infrastructures using fl at the edge. In *2024 IEEE International Conference on Big Data (BigData)* (pp. 4921-4929). IEEE.

Duan, S., Wang, D., Ren, J., Lyu, F., Zhang, Y., Wu, H., & Shen, X. (2022). Distributed artificial intelligence empowered by end-edge-cloud computing: A survey. *IEEE Communications Surveys & Tutorials*, *25*(1), 591-624.

Verma, R. K. (2026). Adaptive AI Models for Real-Time Data Mobility and Security in Urban IoT Networks. In *AI-Based Data Mobility and Intelligent Modeling for Smart Cities* (pp. 37-66). IGI Global Scientific Publishing.

Thallapalli, R., Yadav, B. P., Srinivasulu, B., Venkateshwarlu, P., Srinivas, M., & Kumar, E. K. (2025, October). AI-Driven Cyber-Physical Systems and IoT for Enabling Intelligent, Secure, and Adaptive Ecosystems. In *2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT)* (pp. 1-7). IEEE.

Mohammed, N. U., & Raheem, M. A. R. (2025). Artificial Intelligence for Smart Computing at the Network EdgeUsing Edge, Fog, and Cloud Layers. *Journal of Cognitive Computing and Cybernetic Innovations*, *1*(3), 14-20.

Alkhorasani, H., Zolait, A. H., & Alawag, A. M. (2026, April). Federated Edge Intelligence for Privacy-Preserving IoT-Based Smart Banking Systems: Architectures, Security, and Experimental Challenges. In *2026 International Conference on Frontiers of Engineering and Emerging Technologies (FET)* (pp. 1-6). IEEE.

Rahmati, M., & Pagano, A. (2025, July). Federated learning-driven cybersecurity framework for IoT networks with privacy preserving and real-time threat detection capabilities. In *Informatics* (Vol. 12, No. 3, p. 62). MDPI.

Sindhu, S., Aravindhan, S., Sinha, A., Arvinth, N., & Saidov, N. (2025, December). AI-Driven IoT Security Framework for Real-Time Threat Detection in Large-Scale Smart City Deployments. In *Proceedings of the 9th International Conference on Future Networks and Distributed Systems* (pp. 811-820).

Basani, D. K. R., Gudivaka, R. K., Gudivaka, R. L., Grandhi, S. H., Gudivaka, B. R., & Kamruzzaman, M. M. (2025). Optimizing data privacy and threat detection in cloud-edge collaborative systems: Ai-driven approaches with mcc, pso, and cdns. *International Journal of Automation and Smart Technology*, *15*(1).

Heidari, A., Rastegar, S. H., & Khonsari, A. (2026). Artificial Intelligence-driven privacy preservation in the internet of vehicles: a comprehensive systematic literature review. *Journal of Big Data*, *13*(1), 31.

Paramasivan, M., Sivasubramanian, M., Alagar, K., & Immanuel T, B. (2026). Edge AI and IoT for smart sustainable transportation: Real-time renewable energy optimization–A comprehensive review. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 09544070261464210.

Ranpara, R., Patel, S. K., Kumar, O. P., & Al-Zahrani, F. A. (2025). A computational framework for IoT security integrating deep learning-based semantic algorithms for real-time threat response. *Scientific Reports*, *15*(1), 16794.

Malik, M. T., Muqeet, H. A., Akram, R., Obaid, H. M., Rehman, A. U., & Bermak, A. (2026). Artificial Intelligence Applications in Smart Cities: Integrating Communication, Energy Management, Cybersecurity and Internet of Things. *The Journal of Engineering*, *2026*(1), e70166.