Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.083381

ARTICLE

Quantum-Enhanced Security for Edge-IIoT: Robust Intrusion Detection with a Novel Quantum-Classical Neural Network

Alanoud Al Mazroa[1], Abdulrahman Mohammed Alamoudi[2], Nurdaulet Karabayev[3], Jawad Ahmad[4,*] and Muhammad Shahbaz Khan[5]

[1]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[2]Department of Electrical and Electronic Engineering, College of Engineering, University of Jeddah, Jeddah, Saudi Arabia
[3]Faculty of Information Technology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
[4]Cybersecurity Center, Prince Mohammad Bin Fahd University, Al-Khobar, Saudi Arabia
[5]School of Computer Science and Digital Technologies, Aston University, Birmingham, UK

*Corresponding Author: Jawad Ahmad. Email: jahmad@pmu.edu.sa

Received: 02 April 2026; Accepted: 03 July 2026; Published: 13 August 2026

ABSTRACT: The rapid expansion of the Internet of Things in critical industrial environments has significantly increased the attack surface, exposing systems to sophisticated cyber threats. Traditional pattern-based intrusion detection systems struggle to detect such advanced attacks, while deep learning approaches, despite achieving high detection accuracy, often suffer from high computational cost and latency, limiting their deployment in resource-constrained edge gateways. Quantum machine learning offers a promising alternative by enabling high-dimensional feature representation; however, current implementations are constrained by hardware noise, limited qubit availability and backend-dependent execution characteristics in the Noisy Intermediate-Scale Quantum era. To address these challenges, this paper proposes a Residual Hybrid Quantum-Classical Neural Network (RHQ-CNN) for efficient intrusion detection in Edge-Industrial Internet of Things (Edge-IIoT) environments under idealized quantum simulation conditions. The proposed framework employs a classical neural network encoder to compress high-dimensional network traffic into a compact latent representation suitable for quantum processing. A variational quantum circuit with serial data re-uploading is then utilised to model complex non-linear decision boundaries without increasing qubit requirements. In addition, a residual connection fuses classical and quantum representations to improve training stability and preserve latent feature information. The model is evaluated on the Edge-IIoTset dataset and achieves a test accuracy of 99.94%, with high weighted performance across the 15-class detection task, although the extremely low-sample Fingerprinting class remains comparatively more challenging. Additional controlled ablation, deployment-cost, and bootstrap confidence interval analyses demonstrate the test-set metric stability and computational trade-offs of the proposed architecture. These findings highlight the potential of hybrid quantum-classical models for next-generation cybersecurity in industrial IoT systems.

KEYWORDS: Quantum machine learning; hybrid quantum-classical neural network; intrusion detection system; edge-IIoT; variational quantum circuit; data re-uploading; residual learning; cybersecurity

1 Introduction

The rapid adoption of the Internet of Things (IoT) in industrial environments has significantly improved automation, efficiency, and real-time monitoring capabilities. However, Industrial IoT (IIoT)

systems generate massive volumes of heterogeneous network traffic and expand the overall attack surface, exposing systems to increasingly sophisticated cyber threats [1]. Billions of interconnected devices at the network edge continuously produce high-dimensional data streams, making real-time analysis challenging. Traditional centralized security architectures struggle to process such large-scale data efficiently, leading to delays in threat detection and response [2]. At the same time, modern attack vectors such as ransomware and Distributed Denial of Service (DDoS) attacks are capable of bypassing conventional rule-based intrusion detection systems (IDS), which rely on predefined signatures and static rules [3,4]. IoT environments are also exposed to botnet-driven malicious activities, including DDoS, reconnaissance, information-theft, and key-logging attacks [5]. These challenges highlight the need for intelligent, adaptive, and data-driven security mechanisms capable of operating in dynamic IIoT environments.

Deep Learning (DL) techniques have emerged as powerful tools for intrusion detection due to their ability to learn complex non-linear patterns from large-scale network traffic data. Models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have demonstrated high accuracy in detecting various types of cyber-attacks [6]. Benchmark datasets such as Edge-IIoTset [7] have enabled systematic evaluation of IDS models, showing that deep neural networks can effectively classify diverse attack categories including XSS and SQL injection. Recent approaches also incorporate Generative Adversarial Networks (GANs) to generate synthetic attack samples and improve model robustness [8], while Autoencoders are used for unsupervised anomaly detection in zero-day attack scenarios [9].

Despite these advancements, deploying deep learning models in IIoT environments remains challenging. The computational cost, model-size requirements, and inference latency of deep-learning-based intrusion detection systems can hinder their deployment on resource-constrained IoT edge devices [10]. This creates a trade-off between detection accuracy and deployment feasibility, especially in real-time and low-latency industrial systems. Quantum Machine Learning (QML) has recently emerged as a promising paradigm to address some of the limitations of classical learning methods. QML can encode classical inputs as quantum states, thereby mapping the data into a quantum Hilbert space where quantum kernels or variational quantum models can be used for classification [11]. Hybrid quantum-classical architectures have shown encouraging results, achieving competitive accuracy with fewer trainable parameters compared to purely classical approaches [12].

However, practical deployment of QML is constrained by the limitations of Noisy Intermediate-Scale Quantum (NISQ) hardware. Current quantum devices have limited qubit counts, short coherence times, and are highly sensitive to noise, making it difficult to directly encode high-dimensional IIoT data into quantum circuits [13]. Additionally, deep quantum models suffer from issues such as barren plateaus, where gradients vanish during training, leading to unstable optimisation and reduced learning efficiency [14]. These challenges highlight the need for carefully designed hybrid models that reduce qubit requirements while remaining suitable for simulation-based architectural evaluation under current NISQ-era constraints. In this work, NISQ compatibility refers to the compact four-qubit architectural design rather than experimentally verified deployment on physical quantum hardware.

To address these limitations, this paper proposes a novel Residual Hybrid Quantum-Classical Neural Network (RHQ-CNN) with a data re-uploading strategy. The proposed design targets a compact simulation-based hybrid architecture through three key contributions. First, a lightweight classical encoder is employed to compress IIoT traffic data into a compact latent representation that is suitable for quantum processing, similar in spirit to variational autoencoding approaches [15]. Second, a serial data re-uploading mechanism is adopted, guided by [16], enabling a compact four-qubit circuit to approximate complex decision boundaries without increasing the number of qubits. Third, a classical skip connection is introduced to stabilise training and provide a reliable gradient flow, mitigating optimisation challenges associated with quantum circuits.

The integration of these components allows the quantum module to enhance overall model performance while maintaining efficiency and practicality [17,18].

The remainder of this paper is organised as follows. Section 2 reviews the related work in both classical and quantum-based intrusion detection systems. Section 3 presents the proposed RHQ-CNN framework, including its mathematical formulation and architectural design. Section 4 describes the experimental setup, including dataset details, preprocessing steps, and evaluation methodology. Section 5 discusses the experimental results, including performance analysis and comparative evaluation. Finally, Section 6 concludes the paper and outlines directions for future research.

2 Related Work

Intrusion Detection Systems for IIoT environments have evolved significantly in recent years. Existing research can be broadly categorised into four main directions: (i) classical machine learning-based IDS, (ii) deep learning-based IDS, (iii) quantum machine learning-based IDS, and (iv) hybrid quantum-classical IDS. Each category offers unique advantages but also presents important limitations in terms of scalability, computational efficiency, and practical deployment.

Classical ML techniques have been widely adopted for intrusion detection in IIoT due to their simplicity, interpretability, and relatively low computational cost. Several studies demonstrate that lightweight ML models can achieve strong performance across different datasets. For instance, Ismail et al. [19] conducted a comprehensive comparative study using multiple IIoT datasets, showing that ensemble methods such as Random Forest consistently achieve high accuracy, while Decision Trees offer a good balance between performance and efficiency. Similarly, Awad et al. [20] reported extremely high detection accuracy using Decision Trees on the Edge-IIoTset dataset, highlighting the effectiveness of simple models in controlled settings. Eid et al. [21] further demonstrated that careful preprocessing and feature engineering can significantly improve ML-based IDS performance, achieving near-perfect results using Random Forest and Decision Tree models. In addition, Gaber et al. [22] showed that optimisation-based feature selection (e.g., Particle Swarm Optimisation and Bat Algorithm) can reduce dimensionality and improve detection accuracy while maintaining efficiency. To address edge-IIoT constraints, Yao et al. [23] proposed a hierarchical architecture combining lightweight ML models at the edge with deep learning models at the central layer, achieving a balance between computational efficiency and detection accuracy. Similarly, Tiwari et al. [24] focused on lightweight IDS design for edge deployment, emphasising feature optimisation and low-latency processing. Despite these advancements, ML-based IDS approaches suffer from several limitations. Their performance is often highly dependent on dataset characteristics and preprocessing strategies, and they may struggle to detect complex or evolving attack patterns. Furthermore, most studies rely on offline evaluation and do not fully address real-time deployment challenges in dynamic IIoT environments.

DL approaches have been introduced to overcome the limitations of classical ML by enabling automatic feature extraction and improved detection of complex attack patterns. These models are particularly effective in handling high-dimensional and heterogeneous IIoT data. Soliman et al. [25] proposed a hybrid ML-DL framework incorporating LSTM, Bi-LSTM, and GRU models, achieving extremely high accuracy in both binary and multi-class classification tasks. Similarly, Popoola et al. [26] introduced a multi-stage deep learning architecture that improves detection performance by decomposing the classification process into sequential stages. Lightweight DL models have also been explored to address resource constraints. Mendonça et al. [27] proposed a sparse evolutionary neural network that significantly reduces model complexity while maintaining high accuracy, enabling deployment on low-cost edge devices. To enhance interpretability, Shtayat et al. [28] integrated ensemble CNN models with explainable AI techniques such as SHAP and LIME, improving both detection performance and transparency. Similarly, Nandanwar and Katarya [29]

demonstrated that deep neural networks can effectively capture non-linear relationships in IIoT traffic, leading to improved generalisation. However, DL-based IDS models introduce significant computational overhead and complexity. They often require extensive training time, large datasets, and powerful hardware, making real-time deployment on resource-constrained IIoT devices challenging. In addition, many DL models operate as black boxes, limiting interpretability and trust in security-critical applications.

QML has recently emerged as a promising approach for improving IDS performance, particularly in large-scale and high-dimensional data environments. QML models aim to leverage quantum properties such as superposition and entanglement to enhance learning efficiency. Kalinin and Krundyshev [30] demonstrated that QML models such as QSVM and QCNN can achieve competitive accuracy while reducing training time on large datasets. Similarly, Abreu et al. [12] proposed a hybrid QML-IDS framework compatible with NISQ devices, integrating variational quantum circuits with classical optimisation. More recent work by Bharathi et al. [5] introduced hybrid quantum neural networks (HQNN) with error mitigation techniques, showing improved scalability and performance. Shaji et al. [31] evaluated multiple QML models, including QCNN and quantum random forests, demonstrating promising results under limited training data conditions. However, several critical challenges remain. Most QML-based IDS approaches require heavy preprocessing, including dimensionality reduction, to fit data into limited qubit systems. Cirillo et al. [32] highlighted that QML models are highly sensitive to noise and circuit depth, and current results do not demonstrate a clear practical quantum advantage. Furthermore, many studies rely on simulated environments rather than real quantum hardware. In summary, while QML offers theoretical benefits, its practical deployment is limited by hardware constraints, noise sensitivity, and high implementation complexity.

To address the limitations of fully quantum systems, recent research has focused on hybrid quantum-classical IDS architectures, where quantum components are integrated into classical pipelines. Salek et al. [33] proposed a hybrid framework using a quantum RBM for classification, demonstrating improved performance over classical models. Elsedimy et al. [34] combined QSVM with metaheuristic optimisation, achieving high accuracy on large datasets. Similarly, Kuo et al. [35] used quantum-inspired optimisation techniques to improve feature selection and model tuning. More advanced hybrid approaches have been developed using deep learning architectures. Sahoo et al. [15] integrated quantum embeddings with CNN and BiLSTM models, while Amara et al. [36] proposed a QCNN-based IDS for feature extraction and classification. Despite strong performance, these hybrid approaches share several common limitations. First, they rely on complex multi-stage pipelines involving preprocessing, feature engineering, and optimisation, which increase system complexity and computational cost. Second, the quantum component is typically limited to feature encoding or transformation, while the main learning and decision-making processes remain classical. Third, the contribution of quantum computation is rarely isolated or rigorously validated, making it difficult to demonstrate a clear quantum advantage. Table 1 summarises representative machine learning (ML), DL, QML, and hybrid quantum-classical IDS studies, including their datasets, methods, and reported performance.

Table 1: Comparison of IDS approaches in IIoT across ML, DL, QML, and Hybrid approaches.

Paper	Model/Approach	Dataset(s)	Key Performance/Findings
Machine Learning-Based IDS for IIoT			
Ismail et al. (2025) [19]	DT, RF, k-NN, LR	TON_IoT, WUSTL-IIOT-2021, EdgeIIoTset	RF achieves highest accuracy; DT offers competitive performance with lower computational cost; performance varies across datasets.

(Continued)

Table 1 (continued)

Paper	Model/Approach	Dataset(s)	Key Performance/Findings
Awad et al. (2024) [20]	k-NN, DT, NN	Edge-IIoTset	DT achieves 100% (binary) and 98.46% (multiclass); NN and k-NN ~98% (binary), ~94%–95% (multiclass).
Eid et al. (2023) [21]	RF, DT, KNN, SVM, LR, NB	WUSTL-IIoT-2021	RF achieves 99.97% MCC; DT and KNN > 99%; NB lower (87.42%).
Yao et al. (2019) [23]	LightGBM + DL hybrid	Not specified	Improved accuracy with reduced training time; ~90% bandwidth reduction using feature transmission.
Gaber et al. (2023) [22]	BA/PSO + RF, k-NN, MLP	WUSTL-IIoT-2021	BA+RF: 99.99% accuracy, 0.996 precision/recall/F1; feature selection improves efficiency.
Tiwari et al. (2024) [24]	Lightweight ML + feature optimisation	Not specified	High accuracy with reduced computational overhead; suitable for real-time edge deployment.
Deep Learning-Based IDS in IIoT			
Soliman et al. (2023) [25]	LSTM, Bi-LSTM, GRU + ML	ToN_IoT	Up to 99.999% (binary) and 99.98% (multiclass); DL outperforms ML.
Popoola et al. (2025) [26]	Multi-stage CNN + RNN	Not specified	Multi-stage architecture improves accuracy and reduces false positives.
Mendonça et al. (2022) [27]	SET-based MLP	DS2OS, CICIDS2017	~99% accuracy; 6.25% improvement; testing tim ~2.29 ms.
Shtayat et al. (2023) [28]	Ensemble CNN + SHAP/LIME	ToN_IoT	99.69% (binary), 99.63% (multiclass); improved interpretability.
Nandanwar & Katarya (2024) [29]	DNN	Benchmark IIoT datasets	High accuracy with reduced false positives; improved generalisation.
Quantum Machine Learning for Intrusion Detection			
Kalinin & Krundyshev (2023) [30]	QSVM, QCNN	Custom (10M+ records, IoT-23, BoT-IoT derived)	~98% accuracy; ~2× faster training.

(Continued)

Table 1 (continued)

Paper	Model/Approach	Dataset(s)	Key Performance/Findings
Abreu et al. (2024) [12]	VQC, QSVM, QCNN	UNSW-NB15, CICIDS2017, CICIoT2023	QCNN up to 95.60% F1; performance depends on hardware.
Bharathi et al. (2026) [5]	HQNN	Not specified	Higher accuracy and lower error vs. classical ML; improved scalability.
Salvakkam et al. (2023) [37]	CNN + GRU + RBM + FHE	KDD, NSL-KDD, UNSW-NB15	~97.88% accuracy; improved secure processing.
Cirillo et al. (2026) [32]	QSVC, VQC, HQNN	ToN_IoT, NSL-KDD	QSVC ~94.6%; performance degrades under noise.
Shaji et al. (2025) [31]	QCNN, QSVM, Q-RF, Q-KMeans	ToN-IoT	Q-RF ~95%, QCNN ~94%; requires heavy PCA reduction.
Hybrid Quantum-Classical IDS			
Salek et al. (2023) [33]	NN + Quantum RBM	CAN dataset (vehicle)	97%–98.3% accuracy; better than classical RBM.
Elsedimy et al. (2024) [34]	QSVM + IGWO	BoT-IoT	99.11% accuracy; strong improvement over ML baselines.
Kuo et al. (2024) [35]	DNN + GQTS	CICIDS2017	98.88% (binary), 98.82% (multiclass); 60% feature reduction.
Sahoo et al. (2025) [15]	CNN + BiLSTM + QAOA embedding	CIC-DDoS2019	92.63% accuracy; improved feature representation.
Amara et al. (2025) [36]	QCNN + classical NN	BoT-IoT	~99% accuracy; reduced false negatives vs. CNN.

Summary and Research Gap

From the above discussion, several key observations can be made:

- Classical ML-based IDS are efficient but struggle with complex and evolving attack patterns.
- Deep learning models improve detection performance but introduce high computational overhead and limited interpretability.
- Quantum machine learning offers theoretical advantages but faces significant practical challenges, including noise sensitivity and hardware limitations.
- Hybrid quantum-classical approaches are more practical, but current designs are complex and do not clearly exploit quantum advantages.

Most importantly, existing hybrid IDS frameworks utilise quantum components primarily as feature transformation modules rather than integrating them into the core decision-making process. This results in increased complexity without clear evidence of the specific contribution made by the quantum component.

Therefore, there is a clear need for lightweight and targeted hybrid architectures that:

- minimise preprocessing and architectural complexity,

- integrate quantum computation into critical decision components, and
- provide clearer and more measurable evidence of quantum contribution under controlled experimental conditions.

In this work, the term quantum contribution is used cautiously to describe the measurable effect of specific quantum design components, such as data re-uploading and entanglement, under idealized simulation conditions. This study does not claim definitive practical quantum advantage on physical hardware; rather, it evaluates whether the proposed quantum-classical design improves classification performance compared with carefully matched classical and simplified quantum alternatives.

3 Proposed Methodology

The proposed RHQ-CNN framework is designed to operate within the constraints of NISQ devices while maximizing classification efficacy for high-dimensional IIoT traffic. The architecture consists of three mathematically distinct stages: (i) classical latent feature extraction, (ii) quantum processing using serial data re-uploading, and (iii) residual feature fusion. The comprehensive workflow is illustrated in Fig. 1.

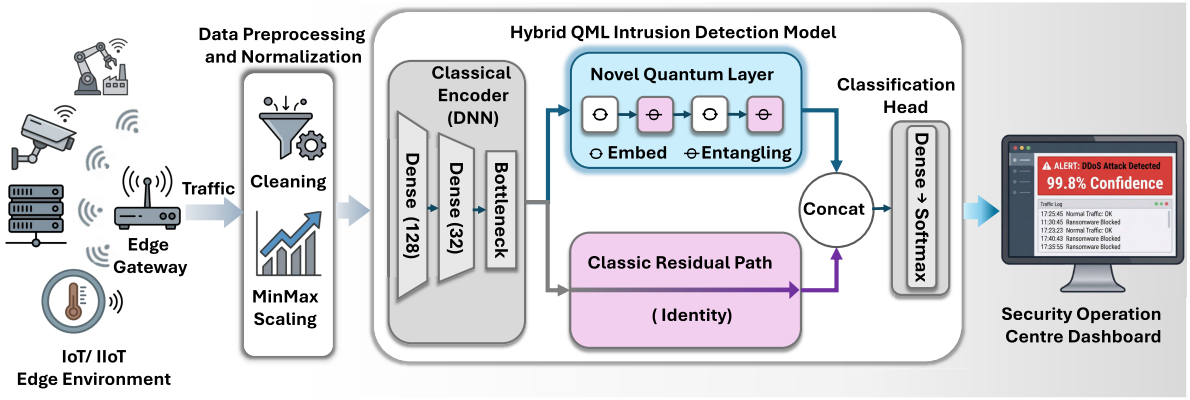

Figure 1: The proposed RHQ-CNN framework for Edge-IIoT intrusion detection. The architecture uses a classical encoder for dimensionality reduction, a VQC with data re-uploading for non-linear feature mapping and a residual path to preserve fidelity.

3.1 Dataset Description and Class Distribution

The model performance is evaluated using the Edge-IIoTset dataset which represents diverse conditions in real-world IoT and IIoT environments [7]. The dataset contains both normal and attack traffic collected from heterogeneous IoT/IIoT devices and communication protocols. It includes one normal class and fourteen attack classes.

Before duplicate and missing-value removal, the dataset contained 2,219,201 labelled records across 15 classes. After preprocessing, the final cleaned dataset contained 1,983,328 samples, as reported in Table 2. Fig. 2 highlights the significant class imbalance within the dataset. Normal traffic remained the majority class with 1,435,546 samples, corresponding to 72.38% of the cleaned dataset. In contrast, minority classes such as Fingerprinting and MITM contained only 856 and 394 samples, corresponding to 0.04% and 0.02%, respectively. This imbalance creates a challenge for classifier performance and makes overall accuracy alone insufficient for evaluation. Therefore, macro-averaged metrics such as Macro Precision, Macro Recall, Macro F1-score, and Macro PR-AUC were used to better assess detection performance across both majority and minority attack classes.

Table 2: Distribution of traffic classes in the Edge-IIoTset dataset before and after duplicate and missing-value removal.

Class Label	Original Samples	Removed Records	Final Samples	Percentage (%)
Normal	1,615,643	180,097	1,435,546	72.38
DDoS (UDP)	121,568	1	121,567	6.13
DDoS (ICMP)	116,436	48,494	67,942	3.43
SQL Injection	51,203	227	50,976	2.57
Password	50,153	74	50,079	2.52
Vulnerability Scanner	50,110	82	50,028	2.52
DDoS (TCP)	50,062	0	50,062	2.52
DDoS (HTTP)	49,911	74	49,837	2.51
Uploading	37,634	398	37,236	1.88
Backdoor	24,862	827	24,035	1.21
Port Scanning	22,564	2576	19,988	1.01
XSS	15,915	830	15,085	0.76
Ransomware	10,925	1228	9697	0.49
Fingerprinting	1001	145	856	0.04
MITM	1214	820	394	0.02
Total	2,219,201	235,873	1,983,328	100.00

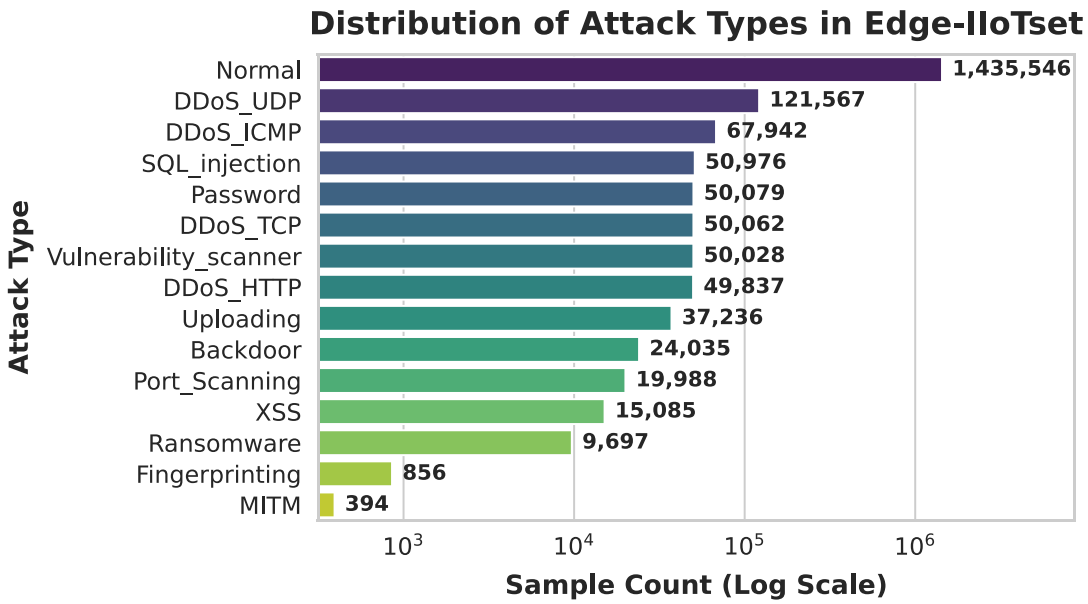

Figure 2: Attack type distribution in the Edge-IIoTset shown on a logarithmic scale. The scale highlights class imbalance where Normal traffic exceeds minority attacks such as Fingerprinting and MITM.

Furthermore, to clarify the attack landscape represented in the Edge-IIoTset dataset, Table 3, adapted from [7], summarises the major threat groups and representative attack types used in the dataset. The Edge-IIoTset dataset provides high-correlation traffic features extracted using Zeek and TShark from multiple network and IoT/IIoT protocol layers. Instead of listing all individual attributes, Table 4 summarises the main feature groups used in this study.

Table 3: Attack categories and representative scenarios in the Edge-IIoTset dataset.

Threat Group	Representative Attacks	Primary Security Impact	Typical Execution Approach
DoS/DDoS	TCP SYN Flood, UDP Flood, HTTP Flood, ICMP Flood	Service disruption, resource exhaustion, and denial of legitimate access	High-volume crafted packets or repeated request flooding against edge services
Information Gathering	Port Scanning, OS Fingerprinting, Vulnerability Scanning	Reconnaissance of hosts, exposed services, operating systems, and weaknesses	Active probing and scanning of networked devices and web-facing components
Man-in-the-Middle	DNS Spoofing, ARP Spoofing	Traffic interception, redirection, and manipulation of communications	Spoofing-based impersonation within local network communication paths
Injection Attacks	Cross-Site Scripting (XSS), SQL Injection, Uploading Attack	Compromise of application integrity, unauthorised access, and malicious code insertion	Injection of malicious input through vulnerable web interfaces and file upload paths
Malware Attacks	Backdoor, Password Cracking, Ransomware	Persistent compromise, credential abuse, and data encryption or takeover	Malware deployment, repeated credential guessing, or post-compromise encryption

Table 4: Feature categories in the Edge-IIoTset dataset used for intrusion detection.

Feature Group	Example Features	Data Source	Role in Detection
TCP Features	tcp.flags, tcp.srcport, tcp.dstport, tcp.options	Network Traffic	Connection behaviour, session control, and flow patterns
UDP Features	udp.port, udp.stream	Network Traffic	Lightweight communication behaviour and anomaly detection
HTTP Features	http.request.method, http.referer, http.uri	Application Layer	Web-based attacks such as XSS and injection detection
MQTT Features	mqtt.topic, mqtt.msg, mqtt.protoname	IoT Protocol	IoT communication patterns and protocol misuse detection
DNS Features	dns.qry.name.len, dns.flags	Network Traffic	Name resolution behaviour and spoofing detection
ICMP Features	icmp.seq, icmp.checksum	Network Traffic	Ping-based attacks and network probing detection
Modbus Features	mbtcp.len, mbtcp.trans_id	Industrial Protocol	IIoT device communication and control anomalies
General Flow Features	packet length, flow duration	Traffic Metadata	Overall traffic statistics and behavioural patterns

3.2 Dataset Preprocessing and Splitting

Before model training, the Edge-IIoTset dataset was cleaned to reduce redundancy and prevent possible information leakage. Three non-predictive columns, namely frame.time, ip.src_host, and ip.dst_host were removed because timestamp and host-address fields may introduce temporal or device-specific bias rather than generalizable attack behaviour. After removing these columns, records containing missing values were discarded, and duplicate rows were removed while retaining the first occurrence. This cleaning step reduced the dataset from 2,219,201 original records to 1,983,328 final samples as shown in Table 2. Following data cleaning, categorical class labels were encoded into integer labels for multi-class classification. The cleaned dataset was divided into training and held-out testing subsets using an 80:20 stratified split. Stratification was applied to preserve the original class distribution across both subsets, which is particularly important because the dataset is highly imbalanced. During training, a portion of the training set was further used as an internal validation set through the validation split of 0.2. The held-out test set was kept completely separate and was used only for final model evaluation.

The input feature matrix was then normalized using min–max scaling. For the Edge-IIoTset, the input dimension after preprocessing is denoted as D, and the number of output classes is $C = 15$. To ensure compatibility with quantum rotation gates, all numerical input features were bounded. A min–max scaler was applied directly to map each numerical feature to the interval $[0, \pi]$:

$$x'_j = \pi \cdot \frac{x_j - \min(x_j)}{\max(x_j) - \min(x_j)}, \quad \forall j \in \{1, \ldots, D\}. \tag{1}$$

This preprocessing step ensures that the input values remain bounded and can be consistently transformed by the classical encoder before quantum embedding.

3.3 Stage 1: Classical Latent Feature Extraction

Directly mapping $D = 58$ features to a quantum circuit would require $N = 58$ qubits, which exceeds the coherent capacity of current NISQ hardware. Therefore, a classical DNN encoder, denoted as $\mathcal{E}$, is employed to compress the high-dimensional input x' into a compact latent vector $z \in \mathbb{R}^{N_q}$, where $N_q = 4$ is the number of available qubits. Let $W_1 \in \mathbb{R}^{h \times D}$ and $W_2 \in \mathbb{R}^{N_q \times h}$ be the weight matrices, and b_1, b_2 be the bias vectors. The latent representation z is computed as:

$$h = \mathrm{ReLU}(W_1 x' + b_1) \tag{2}$$

$$z = \pi \cdot \tanh(W_2 h + b_2) \tag{3}$$

The hyperbolic tangent activation constrains the latent features to $[-1, 1]$, and the factor π rescales them to $[-\pi, \pi]$, ensuring full angular coverage of the Bloch sphere during quantum encoding.

3.4 Stage 2: Variational Quantum Circuit with Data Re-Uploading

The core innovation of the RHQ-CNN is the integration of a VQC utilizing the *Data Re-uploading* strategy. Unlike standard encoding where data is embedded once, re-uploading interleaves data encoding unitaries $S(z)$ with trainable variational unitaries $V(\theta)$.

Let the initial quantum state be $|0\rangle^{\otimes N_q}$. The circuit is composed of L layers. In each layer l, the latent vector z is re-encoded, followed by a trainable unitary evolution. The global unitary operation $U(z, \theta)$ is defined as:

$$U(z, \theta) = V(\theta_L) S(z) \ldots V(\theta_1) S(z) \tag{4}$$

3.4.1 Angle Embedding (Feature Map)

The encoding unitary $S(z)$ maps the classical latent values to the rotation angles of qubits using Pauli-Y rotations. This is defined as:

$$S(z) = \bigotimes_{k=1}^{N_q} R_y(z_k) = \bigotimes_{k=1}^{N_q} \exp\left(-i\frac{z_k}{2}Y\right) \tag{5}$$

where Y is the Pauli-Y operator.

3.4.2 Strongly Entangling Ansatz

To capture complex correlations between features, the trainable unitary $V(\theta_l)$ typically consists of arbitrary single-qubit rotations followed by a ring of CNOT gates to induce entanglement:

$$V(\theta_l) = \left(\prod_{k=1}^{N_q} \mathrm{CNOT}_{k,(k \mod N_q)+1}\right)\left(\bigotimes_{k=1}^{N_q} R(\alpha_k, \beta_k, \gamma_k)\right) \tag{6}$$

3.4.3 Quantum Measurement

After L layers of processing, the quantum information is extracted by measuring the expectation value of the Pauli-Z operator for each qubit. The quantum feature vector $q \in \mathbb{R}^{N_q}$ is given as follows. To define it, let the final quantum state after L layers be

$$|\psi_f\rangle = U(z,\theta)\,|0\rangle^{\otimes N_q}. \tag{7}$$

$$q_k = \langle\psi_f|\,Z_k\,|\psi_f\rangle = \langle 0|\,U^{\leqq}(z,\theta)\,Z_k\,U(z,\theta)\,|0\rangle. \tag{8}$$

3.5 Stage 3: Post-Quantum Projection, Residual Fusion and Classification

Deep quantum circuits often suffer from the "vanishing gradient" problem, commonly referred to as barren plateaus. To mitigate this issue, a residual skip connection is introduced between the classical latent representation and the final classification head. After quantum measurement, the extracted quantum features q are first passed through a lightweight post-quantum dense projection layer to obtain the refined quantum representation q' before fusion.

The post-quantum projected feature vector q' is defined as:

$$q' = \mathrm{ReLU}\left(W_p q + b_p\right), \tag{9}$$

where $W_p \in \mathbb{R}^{N_q \times N_q}$ and $b_p \in \mathbb{R}^{N_q}$ are the weights and bias of the post-quantum dense projection layer. Since $q \in \mathbb{R}^{N_q}$, the projected quantum representation also satisfies $q' \in \mathbb{R}^{N_q}$. In this study, $N_q = 4$, so this layer maps the measured quantum output from 4 features to 4 projected quantum features.

The fused feature vector f is then defined as:

$$f = z \oplus q', \tag{10}$$

where $\oplus$ denotes the concatenation operation. Since $z \in \mathbb{R}^{N_q}$ and $q' \in \mathbb{R}^{N_q}$, the fused representation satisfies $f \in \mathbb{R}^{2N_q}$.

The fused representation is then passed through a dense classification layer with ReLU activation, followed by a Softmax output layer:

$$h_c = \mathrm{ReLU}\left(W_c f + b_c\right), \tag{11}$$

$$\hat{y} = \text{Softmax}\left(W_o h_c + b_o\right). \tag{12}$$

Here, $h_c \in \mathbb{R}^{16}$ denotes the hidden representation of the classification head, $W_c \in \mathbb{R}^{16\times 2N_q}$ and $b_c \in \mathbb{R}^{16}$ are the weights and bias of the dense classification layer, while $W_o \in \mathbb{R}^{C\times 16}$ and $b_o \in \mathbb{R}^{C}$ are the weights and bias of the Softmax output layer. For the Edge-IIoTset multi-class intrusion detection task, $N_q = 4$ and $C = 15$, so the fused vector has dimension $2N_q = 8$ and the output probability vector satisfies $\hat{y} \in \mathbb{R}^{15}$.

Crucially, this residual fusion preserves the high-fidelity latent information from the classical encoder while allowing the classifier to leverage the post-processed quantum-enhanced non-linear projections produced by the variational quantum circuit.

3.6 Optimization Framework

The model is trained end-to-end by minimizing the Sparse Categorical Cross-Entropy loss function $\mathcal{L}$:

$$\mathcal{L}(\Theta) = -\frac{1}{M}\sum_{i=1}^{M}\sum_{c=0}^{C-1}\mathbb{I}(y^{(i)} = c)\log(\hat{y}_c^{(i)}) \tag{13}$$

where $\Theta = \{W_1, b_1, W_2, b_2, \theta, W_p, b_p, W_c, b_c, W_o, b_o\}$ represents the set of all trainable parameters, including the classical encoder parameters, quantum variational parameters, post-quantum dense projection parameters, classification-head parameters, and Softmax output parameters.

3.7 Detailed Network Architecture

The mathematical formulation defines the operation of RHQ-CNN, while Fig. 3 presents the architectural and layer-level details. The architecture follows a hierarchical pipeline with four functional blocks: Classical Encoder, Hybrid Quantum-Classical Core, Feature Fusion, and Classification Head.

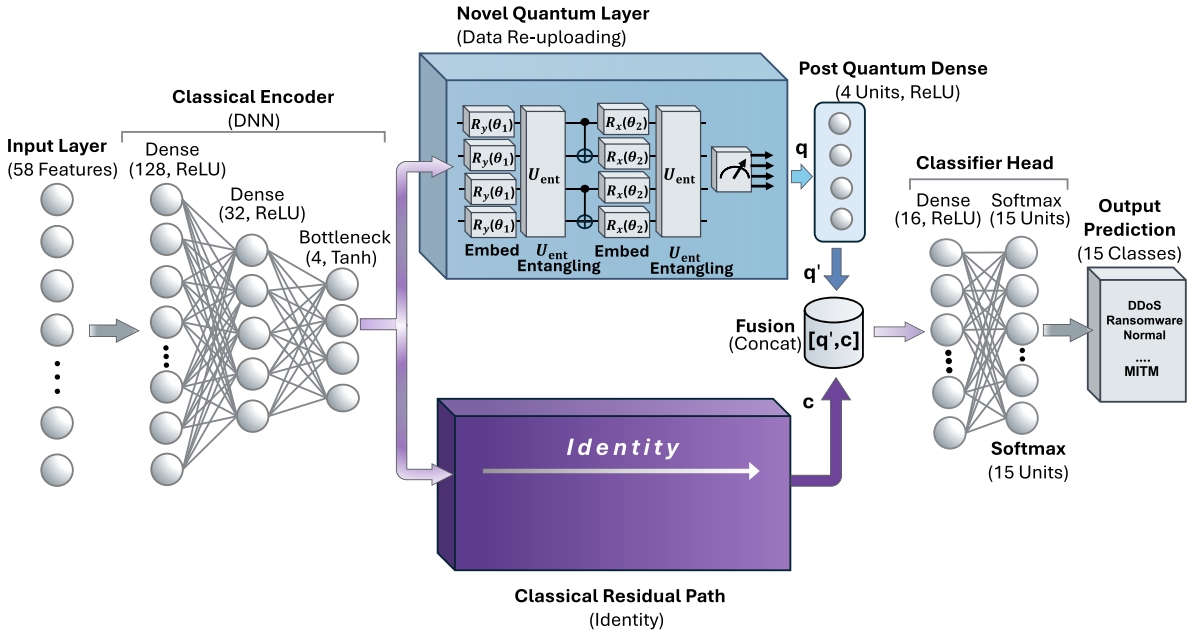

Figure 3: The RHQ-CNN model architecture with details of four distinct functional blocks.

3.7.1 Classical Encoder (Dimensionality Reduction)

The first stage uses a classical DNN to extract high-level features from input vectors with dimension $D = 58$. This stage reduces data dimension to match quantum circuit qubit capacity with $N_q = 4$.

- The input layer maps 58 scaled features into a 128-dimensional space using ReLU activation. This step captures initial linear relations between flow features.
- A second dense layer reduces representation to 32 units and removes noise and irrelevant metadata.
- The final encoder layer maps the 32 features down to a latent vector $z \in \mathbb{R}^4$. A Tanh activation function is applied here to constrain the values to the range $[-1, 1]$. This step enables scaling to rotation angles $[-\pi, \pi]$ for the quantum circuit.

3.7.2 The Hybrid Quantum-Classical Core

At the latent bottleneck, the data flow splits into two parallel processing paths, forming the "Residual Hybrid" structure:

- The latent vector z is fed into a 4-qubit Variational Quantum Circuit. The classical values are encoded as rotation angles on qubits initialized in the $|0\rangle$ state. A sequence of CNOT gates and trainable rotations creates quantum entanglement which allows the circuit to model complex and non-linear dependencies.
- Simultaneously, the original latent vector z bypasses the quantum circuit via an identity skip connection. This path preserves the exact values of the classical features to ensure that the gradient flow remains stable even if the quantum circuit enters a state of vanishing gradients.

3.7.3 Feature Fusion and Classification

The measured quantum output q is first passed through a post-quantum dense projection layer with 4 units, producing the projected quantum feature vector q'. This layer refines the measured quantum representation before fusion while keeping the quantum branch output dimension unchanged. The projected quantum features q' are then concatenated with the 4 classical latent features from the residual identity path, resulting in a combined feature vector of dimension 8. This fused vector is passed to the final classification head, where a dense layer with ReLU activation and a Softmax output layer maps the features to the 15 output classes, generating the final probability distribution for attack detection.

4 Results and Discussion

This section evaluates the proposed RHQ-CNN. Performance is measured using standard classification metrics and confusion matrix analysis. An ablation study and a comparison with state-of-the-art methods are used to benchmark the proposed framework.

4.1 Experimental Setup and Training Protocol

All models were trained and evaluated using the same cleaned dataset, preprocessing pipeline, feature representation, and held-out test set to ensure a fair comparison. The models were trained using the Adam optimizer with sparse categorical cross-entropy loss. During training, a validation split was used from the training data to monitor model performance, while the held-out test set remained unseen until final evaluation. The best-performing model weights were saved according to validation performance and then reloaded for final testing.

The proposed RHQ-CNN uses a classical encoder to compress the preprocessed input features into a four-dimensional latent representation. This latent vector is then supplied to a four-qubit variational

quantum circuit using angle embedding, trainable rotations, entangling gates, and serial data re-uploading. The measured quantum output is first refined using a post-quantum Dense(4) projection layer and is then concatenated with the classical latent vector through a residual connection before being passed to the final Softmax classifier for 15-class intrusion detection. Fixed random seed values were used where supported by NumPy, TensorFlow, and Scikit-learn to improve reproducibility. The implementation was developed using Python with TensorFlow/Keras, PennyLane, NumPy, Pandas, Scikit-learn, XGBoost, and standard scientific computing libraries.

4.2 Overall Classification Performance

The model was evaluated on the held-out test set containing 396,666 samples. Fig. 4 shows the confusion matrix for the 15-class classification task. The RHQ-CNN shows strong diagonal patterns, indicating accurate classification across diverse attack categories. Quantitative results further demonstrate that the model handles class imbalance effectively. Table 5 reports the precision, recall, and F1-score for each class.

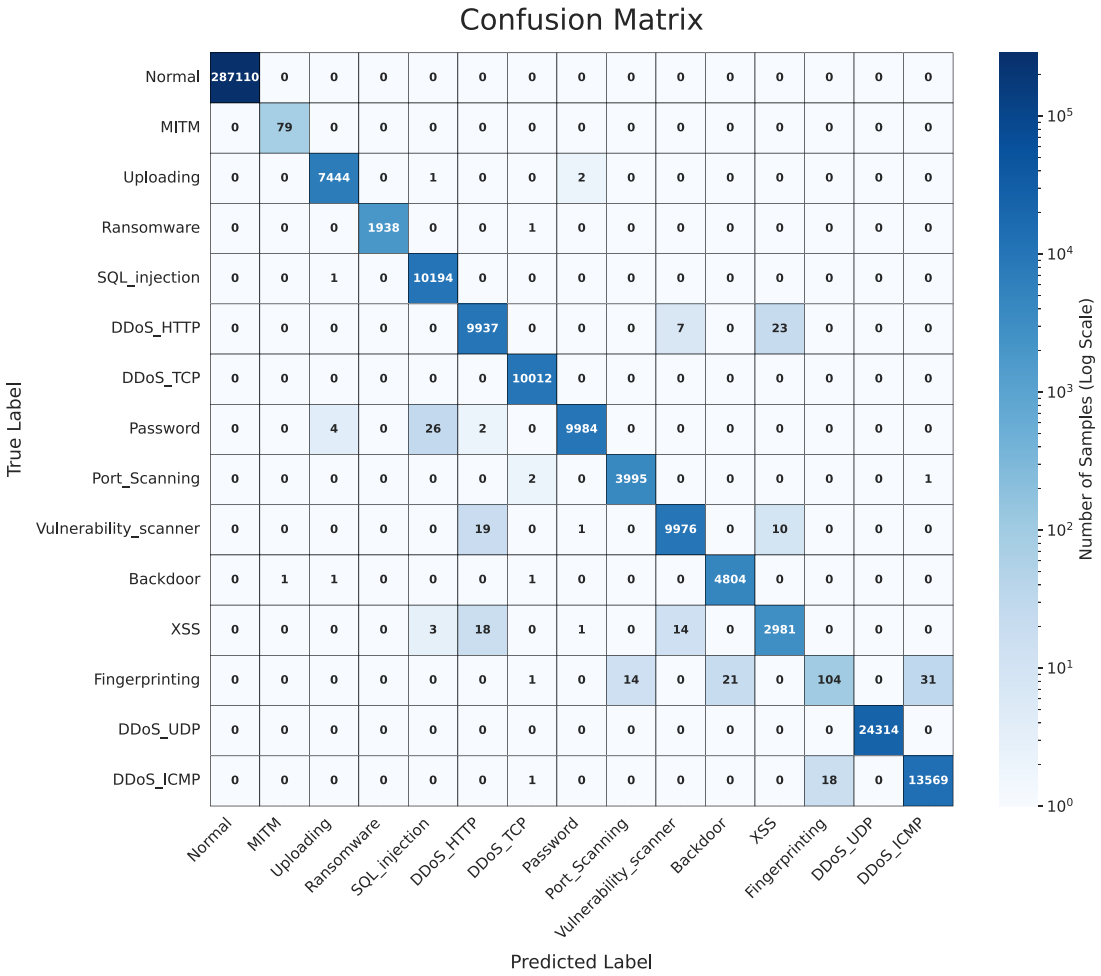

Figure 4: Confusion Matrix of the RHQ-CNN on the Edge-IIoT test set. The matrix shows very low misclassification for majority classes and high precision for minority classes.

Table 5: Detailed classification report for the RHQ-CNN on the Edge-IIoT test set.

Class Label	Precision	Recall	F1-Score	Support
Normal	1.0000	1.0000	1.0000	287,110
DDoS (UDP)	1.0000	1.0000	1.0000	24,314
DDoS (ICMP)	0.9976	0.9986	0.9981	13,588
SQL Injection	0.9971	0.9999	0.9985	10,195
Password	0.9996	0.9968	0.9982	10,016
Vulnerability Scanner	0.9979	0.9970	0.9975	10,006
DDoS (TCP)	0.9994	1.0000	0.9997	10,012
DDoS (HTTP)	0.9961	0.9970	0.9965	9967
Uploading	0.9992	0.9996	0.9994	7447
Backdoor	0.9956	0.9994	0.9975	4807
Port Scanning	0.9965	0.9992	0.9979	3998
XSS	0.9891	0.9881	0.9886	3017
Ransomware	1.0000	0.9995	0.9997	1939
Fingerprinting	0.8525	0.6082	0.7099	171
MITM	0.9875	1.0000	0.9937	79
Accuracy			0.9994	396,666
Macro Average	0.9872	0.9722	0.9783	396,666
Weighted Average	0.9994	0.9994	0.9994	396,666

The model achieves an overall test accuracy of 99.94% on 396,666 held-out test samples. For complex payload attacks, the model achieves an F1-score of 0.9985 for SQL Injection and 0.9886 for XSS. For ultra-minority classes, RHQ-CNN achieves a recall of 1.0000 and an F1-score of 0.9937 for MITM with only 79 test samples. The Fingerprinting class remains more challenging, with a precision of 0.8525, recall of 0.6082, and F1-score of 0.7099 based on 171 test samples. These results indicate stable detection performance even under severe class imbalance.

4.3 Ablation Study: Architectural Impact

An ablation study was conducted to evaluate the contribution of the main architectural components of the proposed RHQ-CNN. The compared variants include a fully classical Simple DNN, a classical residual bottleneck, an RHQ-CNN variant without data re-uploading and entanglement, and the proposed full RHQ-CNN. All models were trained and evaluated using the same cleaned dataset, preprocessing pipeline, feature representation, train-test split, and evaluation protocol to ensure a fair comparison. The Simple DNN was used as a fully classical baseline without residual fusion or quantum processing. The classical residual bottleneck replaced the quantum circuit with a compact classical nonlinear block while preserving the residual structure. The total parameter count of the proposed RHQ-CNN is 12,279 because the non-quantum layers contribute 12,231 parameters, while the PennyLane quantum layer contributes 48 trainable variational parameters. The classical residual bottleneck was parameter-matched to the proposed RHQ-CNN by replacing the PennyLane quantum layer with a compact classical nonlinear bottleneck containing the same 48 trainable parameters. Therefore, both the classical residual bottleneck and the proposed RHQ-CNN contain 12,279 trainable parameters, while the simplified RHQ-CNN contains fewer parameters because removing data re-uploading and entanglement reduces the number of trainable quantum parameters. In this implementation, the variational quantum circuit uses $L = 4$ trainable layers. Since each layer applies

three trainable rotation parameters to each of the $N_q = 4$ qubits, the quantum layer contains $L \times N_q \times 3 = 4 \times 4 \times 3 = 48$ trainable parameters.

The simplified RHQ-CNN retained the residual hybrid structure but removed data re-uploading and quantum entanglement, using only one-time angle encoding and trainable single-qubit rotations. Finally, the proposed RHQ-CNN combined residual fusion, serial data re-uploading, and entangling quantum operations.

As shown in Table 6, the Simple DNN achieved 98.67% accuracy and 93.45% Macro F1-score. The classical residual bottleneck achieved 96.72% accuracy and 84.83% Macro F1-score, while the RHQ-CNN without data re-uploading and entanglement achieved 97.13% accuracy and 85.94% Macro F1-score. In contrast, the proposed RHQ-CNN achieved the best overall performance, with 99.94% accuracy and 97.83% Macro F1-score. These results indicate that the final performance improvement is not due only to the classical encoder or residual path. Instead, the combination of residual fusion, data re-uploading, and quantum entanglement provides a measurable contribution to classification performance under simulation conditions. However, this finding should be interpreted as evidence of a controlled simulation-based quantum contribution rather than a definitive claim of practical quantum advantage on physical quantum hardware.

Table 6: Controlled ablation study isolating the contribution of residual and quantum components.

Model Variant	Parameters	Residual Path	Data Re-Uploading	Quantum Entanglement	Accuracy (%)	Macro F1-Score (%)
Simple DNN	18,383	No	No	No	98.67	93.45
Classical residual bottleneck	12,279	Yes	No	No	96.72	84.83
RHQ-CNN without data re-uploading and entanglement	12,243	Yes	No	No	97.13	85.94
Proposed RHQ-CNN	12,279	Yes	Yes	Yes	99.94	97.83

4.4 Deployment Cost and Edge Feasibility Analysis

To evaluate deployment feasibility, the proposed RHQ-CNN was compared with classical baseline models using the same cleaned dataset, preprocessing pipeline, feature representation, and held-out test set. The comparison includes parameter count, model size, training time, inference latency, throughput, peak RAM usage, accuracy, and Macro F1-score, as shown in Table 7.

Table 7 shows that RHQ-CNN has a compact architecture with 12,279 trainable parameters and a model size of 0.18 MB. It achieves the highest Macro F1-score of 97.83%, outperforming the Classical DNN, XGBoost, and Lightweight 1D-CNN. However, the quantum layer introduces additional simulation overhead, resulting in a higher training time per epoch of 125.26 s and inference latency of 0.0846 ms/sample compared with purely classical baselines. These results indicate that RHQ-CNN is lightweight in terms of storage and parameter count, but its current simulation-based implementation is slower during inference than classical models. Therefore, RHQ-CNN is more suitable for monitoring-oriented Edge-IIoT IDS scenarios where millisecond-level latency is acceptable, while strict ultra-low-latency industrial control systems require further optimization, hardware-aware acceleration, and real quantum hardware validation.

Table 7: Computational and deployment-cost comparison on the Edge-IIoTset dataset.

Model	Parameters	Model Size (MB)	Avg. Train Time/Epoch (s)	Total Training Time (min)	Inference Latency (ms/sample)	Throughput (samples/s)	Peak RAM (MB)	Accuracy (%)	Macro F1 (%)
Classical DNN	18,383	0.24	69.29	11.56	0.0350	28,593.89	3826.66	98.67	93.45
XGBoost	N/A	1.90	N/A	5.58	0.0221	45,151.52	3656.29	98.85	96.49
Lightweight 1D-CNN	64,719	0.78	76.45	12.75	0.0374	26,731.75	4117.40	99.41	95.62
Proposed RHQ-CNN	12,279	0.18	125.26	20.88	0.0846	11,818.44	4121.44	99.94	97.83

4.5 Statistical Reliability under Class Imbalance

The Edge-IIoTset dataset is highly imbalanced, point-estimate metrics alone may not fully demonstrate the reliability of minority-class detection. Therefore, statistical reliability was evaluated using non-parametric bootstrap resampling on the held-out test predictions without retraining the model. In each bootstrap iteration, test samples were resampled with replacement within each class to preserve the original test-set class distribution, and Accuracy, Macro F1-score, and Macro PR-AUC were recomputed. The 2.5th and 97.5th percentiles of the bootstrap distribution were then used to estimate the 95% confidence interval for each metric.

As shown in Table 8, the proposed RHQ-CNN maintains narrow confidence intervals across all reported metrics. The model achieves 99.94% Accuracy with a 95% CI of 99.94%–99.95%, 97.83% Macro F1-score with a 95% CI of 97.38%–98.23%, and 99.18% Macro PR-AUC with a 95% CI of 98.94%–99.39%. These results confirm that the reported performance is statistically stable on the held-out test set and is not solely driven by the dominant Normal traffic class. The narrow confidence intervals indicate low variation across bootstrap samples, supporting the stability of the proposed model under the highly imbalanced test distribution. In particular, the high Macro PR-AUC confirms that RHQ-CNN maintains reliable precision–recall behaviour across minority as well as majority attack classes.

Table 8: Test-set metric stability and confidence interval analysis for RHQ-CNN.

Model	Metric	Point Estimate (%)	95% CI Lower (%)	95% CI Upper (%)
RHQ-CNN	Accuracy	99.94	99.94	99.95
	Macro F1-score	97.83	97.38	98.23
	Macro PR-AUC	99.18	98.94	99.39

It is important to note that this bootstrap analysis was performed without retraining the model. Therefore, the reported confidence intervals estimate test-set metric stability under the held-out prediction distribution, rather than run-to-run stability across different random seeds, parameter initializations, or optimization trajectories. Assessing run-to-run variability would require repeated model training under multiple random seeds, which is left for future work.

4.6 Comparison with State-of-the-Art Studies

As shown in Table 9, the proposed RHQ-CNN achieves competitive performance compared with recent state-of-the-art Edge-IIoTset studies, while using a compact 58-feature input representation and a four-qubit hybrid quantum-classical architecture. Although some recent deep learning models report slightly

higher best-case accuracy, the proposed model provides a compact simulation-validated quantum-classical alternative with strong weighted detection performance. Classical DNN and CNN models achieve baseline accuracy of 94 to 96 percent [7] but show weak performance on minority classes. Recent models such as CST AFNet [38] and SA DCNN [39] achieve accuracy above 99.9%. However, the RHQ-CNN achieves accuracy of 99.94% and matches the performance of complex models while using a compact quantum latent feature space. This performance of the proposed model indicates that quantum-enhanced non linear mapping improves separation of highly correlated attack patterns.

Table 9: Performance comparison with recent state-of-the-art models on Edge-IIoTset.

Reference	Methodology	Classification Setting	No. of Features	Accuracy (%) Best Reported	Reported F1-Score/Weighted F1 (%)
[7]	Deep Neural Network (DNN)	Multi-class	61	94.67	77.33
[39]	SA-DCNN (Self-Attention CNN)	Edge-IIoTset sub-category	29	99.95	99.53
[40]	Autoencoder + Decision Tree	Multi-class	61	99.94	99.94
[38]	CST-AFNet (CNN + BiGRU)	Multi-class	63	99.97	>99.30
[41]	MAGRU	Multi-class	31	99.97	99.64
The Proposed	RHQ-CNN	Multi-class	58	99.94	99.94

The comparison is reported as a literature-level benchmark because prior studies use different preprocessing pipelines, feature sets, class groupings, and train-test splits. The proposed RHQ-CNN is evaluated in a 15-class multi-class setting using 58 features after removing timestamp and host-address identifiers. For the proposed model, the F1-score reported in this literature-level comparison corresponds to the weighted-average F1-score for consistency with commonly reported benchmark results.

4.7 Explainable Artificial Intelligence Analysis

The study uses SHapley Additive exPlanations (SHAP) to improve transparency of the intrusion detection framework and examine how individual network features affect model predictions and decision behavior. SHAP gives a clear method to interpret machine learning models by assigning importance values to input features based on effect on prediction outcomes. The SHAP summary plot is shown in Fig. 5, illustrating the global importance of network features based on the mean absolute SHAP values.

Further, to investigate the explanation patterns learned by the model, SHAP explanation vectors were projected into a two-dimensional space using t-distributed Stochastic Neighbor Embedding (t-SNE). This visual presentation highlights the similarity between different contributions of features.

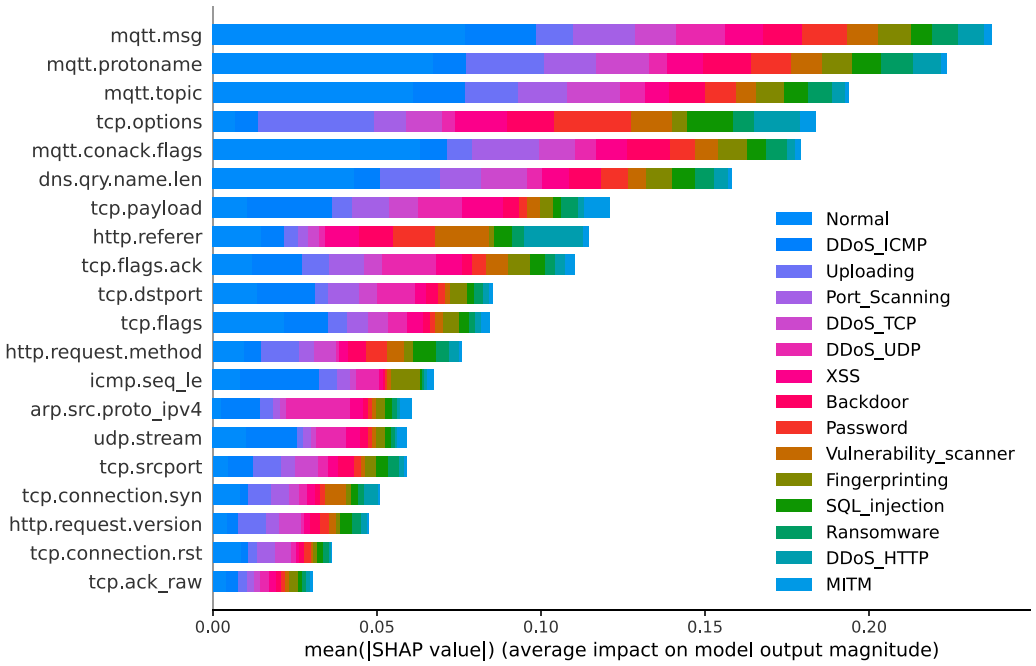

Figure 5: SHAP summary plot illustrating the global importance of network features. Each point represents the SHAP value of a feature for a particular instance, while color indicates the magnitude of the feature value.

Each network traffic instance is shown as a point in Fig. 6 based on its SHAP explanation vector. Instances with similar feature contribution patterns appear closer in the embedded space. The different clusters for attack categories show that the model learns clear explanation patterns for different types of intrusion. Malicious traffic forms compact clusters while normal traffic spreads across a wider region which reflects diverse behavior patterns in IIoT networks.

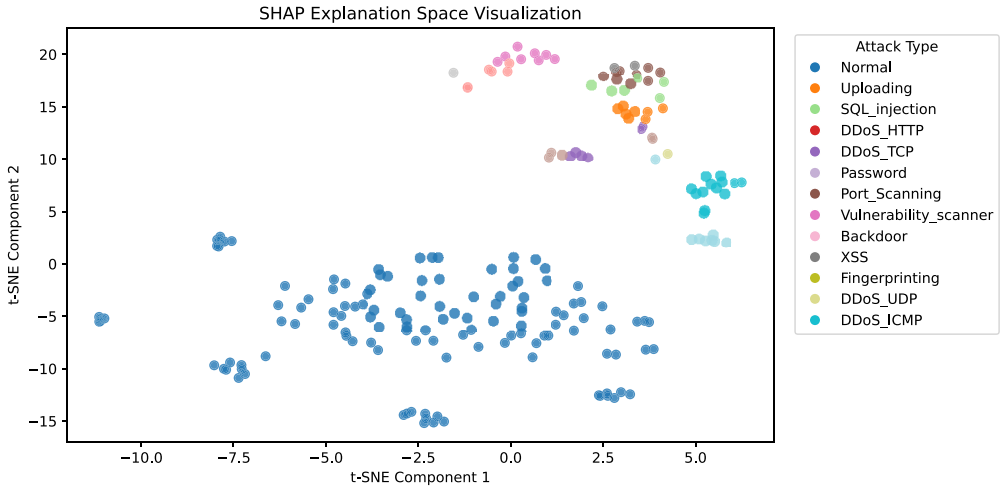

Figure 6: SHAP explanation space visualization using t-SNE. The clear clusters indicate that the model learns distinct explanation patterns for different attack types.

5 Discussion, Limitations, and Future Work

This section discusses the experimental findings and positions RHQ-CNN within the wider IIoT security context. The results show strong detection performance, but important limitations remain. These limitations define the scope of the current study and point toward future research in quantum cybersecurity.

5.1 Discussion of Findings

The results show that the proposed RHQ-CNN achieves strong intrusion detection performance under idealized quantum simulation conditions. The improvement is associated with the combined effect of three architectural components: classical latent compression, residual quantum-classical fusion, and quantum feature transformation through data re-uploading and entanglement.

The classical encoder reduces the high-dimensional Edge-IIoT feature space into a compact four-dimensional latent representation, making the input suitable for four-qubit quantum processing. The residual path preserves this classical latent information and supports stable gradient flow during training. This is important because variational quantum circuits may suffer from optimization instability when the quantum branch alone is used for classification.

The controlled ablation results further indicate that residual learning alone is not sufficient to explain the final performance gain. The classical residual bottleneck and the simplified RHQ-CNN without data re-uploading and entanglement achieved lower Macro F1-scores than the full RHQ-CNN. This suggests that serial data re-uploading and entangling operations provide a measurable contribution to the non-linear feature transformation learned by the model. However, this improvement should be interpreted as simulation-based evidence of quantum contribution rather than a definitive claim of practical quantum advantage on physical quantum hardware. The deployment-cost analysis also shows that RHQ-CNN has a compact parameter count and model size, but its simulation-based quantum layer introduces additional inference overhead compared with purely classical baselines. Therefore, the model is more appropriate for monitoring-oriented Edge-IIoT IDS scenarios where millisecond-level latency is acceptable, while strict real-time industrial control settings require further hardware-aware optimization and real-device validation.

5.2 Limitations of the Study

Although the proposed RHQ-CNN achieves strong simulation-based performance, several limitations must be acknowledged. First, all quantum experiments in this study were conducted using an idealized noiseless quantum simulator. Therefore, the reported results should be interpreted as architectural validation under controlled simulation conditions rather than evidence of hardware-validated NISQ deployment. Real quantum processors are affected by gate infidelity, readout errors, state preparation errors, limited qubit connectivity, calibration drift, and decoherence. These hardware-dependent effects can vary significantly across different quantum backends and even across different calibration cycles of the same backend.

Second, realistic hardware noise behavior was not experimentally evaluated in this work. Generic noise models can provide useful preliminary insight, but they may not fully represent the dynamic and backend-specific behavior of current IBM Q, Rigetti, or similar NISQ-class processors. Hardware performance is also affected by transpilation strategy, qubit mapping, circuit depth, gate scheduling, and temporal drift in device calibration. Therefore, robustness to quantum noise remains an open issue requiring backend-aware evaluation.

Third, although RHQ-CNN has a compact parameter count and small model size, the quantum simulation layer introduces higher inference overhead than purely classical models. The measured latency

may be acceptable for monitoring-oriented IDS scenarios, but strict industrial control loops requiring ultra-low latency may not be feasible without further optimization or dedicated quantum/accelerated hardware support. Thus, RHQ-CNN should not be considered a fully optimized real-time edge deployment solution at this stage.

Finally, the proposed model compresses the input feature space into a four-qubit latent representation. While this design is suitable for compact simulation and small-qubit experiments, scaling the model to larger latent spaces may increase circuit depth, optimization difficulty, and noise sensitivity. Future work must therefore examine the trade-off between qubit count, circuit expressivity, latency, and noise robustness.

5.3 Future Work

Future work will focus on hardware-aware and noise-aware validation of the proposed RHQ-CNN. First, backend-specific noise models will be used to evaluate the effects of gate error, readout error, decoherence, and limited qubit connectivity on classification performance. Second, backend-aware transpilation and qubit mapping will be investigated to reduce circuit depth and improve execution reliability on physical quantum processors.

Third, noise-aware training and error mitigation strategies, such as measurement error mitigation and zero-noise extrapolation, will be explored to improve robustness under realistic NISQ conditions. In addition, the model will be evaluated on real quantum hardware using representative subsets of Edge-IIoT traffic to quantify the gap between ideal simulation performance and hardware-executed performance.

Future studies may also extend RHQ-CNN toward federated and privacy-preserving IIoT intrusion detection, where multiple edge gateways can collaboratively train detection models without sharing raw traffic data. This direction is particularly relevant for distributed industrial environments where privacy, latency, and security constraints must be considered together.

6 Conclusion

This study addressed the critical task of securing the IIoT against advanced and rapidly evolving cyber intrusions. The RHQ-CNN intrusion detection framework was proposed and evaluated using the Edge-IIoTset dataset under an idealized quantum simulation setting with realistic IIoT traffic traces. The main contribution of this work was the integration of classical deep learning with a compact variational quantum circuit. The architecture employed a classical encoder for dimensionality reduction and a quantum circuit with serial data re-uploading for non-linear feature transformation. Furthermore, a classical residual skip connection was introduced to preserve latent feature information and support stable training. RHQ-CNN achieved a test accuracy of 99.94% and a Macro F1-score of 97.83%, demonstrating strong overall classification performance under a highly imbalanced 15-class intrusion detection setting. The model achieved an F1-score of 0.9985 for SQL Injection and 0.9886 for XSS. For ultra-minority classes, RHQ-CNN achieved a recall of 1.0000 and an F1-score of 0.9937 for MITM. However, the Fingerprinting class remained comparatively more challenging, with a recall of 0.6082 and an F1-score of 0.7099. These results indicate that the proposed model maintains high overall detection performance, while also highlighting the need for further improvement in extremely low-sample attack categories. The controlled ablation study further confirmed the contribution of the proposed architectural components. The Simple DNN achieved 98.67% accuracy and 93.45% Macro F1-score, the classical residual bottleneck achieved 96.72% accuracy and 84.83% Macro F1-score, and the RHQ-CNN variant without data re-uploading and entanglement achieved 97.13% accuracy and 85.94% Macro F1-score. In contrast, the full RHQ-CNN achieved the best overall performance among the evaluated variants, with 99.94% accuracy and 97.83% Macro F1-score. This indicates that the combination of residual fusion, serial data re-uploading, and quantum entanglement provides a measurable contribution

to classification performance under controlled simulation conditions. In summary, the proposed RHQ-CNN demonstrates strong potential as a simulation-validated hybrid quantum-classical IDS architecture for Edge-IIoT environments. However, the current study does not claim hardware-proven quantum advantage or deployment-ready NISQ robustness. The experiments were conducted under idealized quantum simulation conditions, and real quantum devices may introduce backend-dependent noise, calibration drift, limited connectivity, and decoherence. Therefore, future work will focus on backend-aware noise modelling, noise-aware training, error mitigation, improved handling of ultra-minority attack classes, and validation on real quantum processors to assess the practical feasibility of RHQ-CNN in real-world IIoT security environments.

Acknowledgement: Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, through the Researchers Supporting Project number (PNURSP2026R510).

Funding Statement: Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, through the Researchers Supporting Project number (PNURSP2026R510).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Muhammad Shahbaz Khan; methodology, Muhammad Shahbaz Khan and Alanoud Al Mazroa; software, Muhammad Shahbaz Khan and Abdulrahman Mohammed Alamoudi; validation, Jawad Ahmad and Nurdaulet Karabayev; formal analysis, Jawad Ahmad and Nurdaulet Karabayev; investigation, Alanoud Al Mazroa and Abdulrahman Mohammed Alamoudi; writing—original draft preparation, Muhammad Shahbaz Khan and Alanoud Al Mazroa; writing—review and editing, Abdulrahman Mohammed Alamoudi, Nurdaulet Karabayev and Jawad Ahmad; visualization, Muhammad Shahbaz Khan; supervision, Muhammad Shahbaz Khan; project administration, Jawad Ahmad; funding acquisition, Alanoud Al Mazroa and Abdulrahman Mohammed Alamoudi. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are openly available in IEEE Dataport at https://dx.doi.org/10.21227/mbc1-1h68.

Ethics Approval: Not applicable.

Conflicts of Interest: Given his role as Editorial Board Member of this journal, Jawad Ahmad had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. The authors declare no other conflicts of interest.

References

1. Mekala SH, Baig Z, Anwar A, Zeadally S. Cybersecurity for Industrial IoT (IIoT): threats, countermeasures, challenges and future directions. Comput Commun. 2023;208:294–320.
2. Fazeldehkordi E, Grønli TM. A survey of security architectures for edge computing-based IoT. Internet Things. 2022;3(3):332–65. doi:10.3390/iot3030019.
3. Slevi ST, Visalakshi P. A survey on deep learning based intrusion detection systems on Internet of Things. In: Proceedings of the 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC); 2021 Nov 11–13; Palladam, India. p. 1488–96.
4. Bansal K, Singhrova A. Review on intrusion detection system for IoT/IIoT-brief study. Multimed Tools Appl. 2024;83(8):23083–108. doi:10.1007/s11042-023-16395-6.
5. Bharathi I, Sonai V, Sridevi S. Quantum-driven enhanced machine learning algorithm for intrusion detection in internet of things environment. EPJ Quantum Technology. 2026;13:20.
6. Naji M, Zougagh H. Deep learning models for cybersecurity in IoT networks. In: Business Intelligence (CBI 2023). Berlin/Heidelberg, Germany: Springer; 2023. p. 29–43.
7. Ferrag MA, Friha O, Hamouda D, Maglaras L, Janicke H. Edge-IIoTset: a new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning. IEEE Access. 2022;10:40281–306.

8. Rahman MA, Francia GA, Shahriar H. Leveraging GANs for synthetic data generation to improve intrusion detection systems. J Future Artif Intell Technol. 2025;1(4):429–39. doi:10.62411/faith.3048-3719-52.
9. Long C, Xiao J, Wei J, Zhao J, Wan W, Du G. Autoencoder ensembles for network intrusion detection. In: Proceedings of the 2022 24th International Conference on Advanced Communication Technology (ICACT); 2022 Feb 13–16; Online. p. 323–33.
10. Hosseininoorbin S, Layeghy S, Sarhan M, Jurdak R, Portmann M. Exploring edge TPU for network intrusion detection in IoT. J Parallel Distr Comput. 2023;179(1):104712. doi:10.1016/j.jpdc.2023.05.001.
11. Schuld M, Killoran N. Quantum machine learning in feature hilbert spaces. Phys Rev Lett. 2019;122(4):040504. doi:10.1103/physrevlett.122.040504.
12. Abreu D, Rothenberg CE, Abelém A. QML-IDS: quantum machine learning intrusion detection system. In: Proceedings of the 2024 IEEE Symposium on Computers and Communications (ISCC); 2024 Jun 26–29; Paris, France. p. 1–6.
13. Keçeci M. Accuracy, noise, and scalability in quantum computation: strategies for the NISQ era and beyond. Open Sci Artic. 2025;2(1):15515113.
14. Qi H, Wang L, Zhu H, Gani A, Gong C. The barren plateaus of quantum neural networks: review, taxonomy and trends. Quantum Inf Process. 2023;22(12):435. doi:10.1007/s11128-023-04188-7.
15. Sahoo JP, Kar B, Chen F. QCShield: a quantum-classical hybrid framework for intrusion detection in edge-IIoT. In: Advances in distributed computing and machine learning (ICADCML 2025). Berlin/Heidelberg, Germany: Springer; 2025. p. 1–18.
16. Fan L, Situ H. Compact data encoding for data re-uploading quantum classifier. Quantum Inf Process. 2022;21(3):87. doi:10.1007/s11128-022-03429-5.
17. Verma M, Arora K, Kumar P, Trivedi K, Yadav J, Pandey SK. Quantum convolutional neural networks with residual learning: advancing quantum model architectures. In: Computer vision, pattern recognition, image processing, and graphics (NCVPRIPG 2025). Berlin/Heidelberg, Germany: Springer; 2025. p. 450–64.
18. Kashif M, Shafique M. ResQuNNs: towards enabling deep learning in quantum convolution neural networks. arXiv:2402.09146. 2024.
19. Ismail S, Dandan S, Qushou A. Intrusion detection in IoT and IIoT: comparing lightweight machine learning techniques using TON_IoT, WUSTL-IIOT-2021, and EdgeIIoTset datasets. IEEE Access. 2025;13:73468–85.
20. Awad OF, Hazim LR, Jasim AA, Ata O. Enhancing IIoT security with machine learning and deep learning for intrusion detection. Malays J Comput Sci. 2024;37(2):139–53. doi:10.22452/mjcs.vol37no2.3.
21. Eid AM, Nassif AB, Soudan B, Injadat MN. IIoT network intrusion detection using machine learning. In: Proceedings of the 2023 6th International Conference on Intelligent Robotics and Control Engineering (IRCE); 2023 Aug 4–6; Jilin, China. p. 196–201.
22. Gaber T, Awotunde JB, Folorunso SO, Ajagbe SA, Eldesouky E. Industrial internet of things intrusion detection method using machine learning and optimization techniques. Wirel Commun Mob Comput. 2023;2023(1):3939895. doi:10.1155/2023/3939895.
23. Yao H, Gao P, Zhang P, Wang J, Jiang C, Lu L. Hybrid intrusion detection system for edge-based IIoT relying on machine-learning-aided detection. IEEE Netw. 2019;33(5):75–81. doi:10.1109/mnet.001.1800479.
24. Tiwari RS, Lakshmi D, Das TK, Tripathy AK, Li KC. A lightweight optimized intrusion detection system using machine learning for edge-based IIoT security. Telecommun Syst. 2024;87(3):605–24. doi:10.1007/s11235-024-01200-y.
25. Soliman S, Oudah W, Aljuhani A. Deep learning-based intrusion detection approach for securing industrial Internet of Things. Alex Eng J. 2023;81:371–83. doi:10.1016/j.aej.2023.09.023.
26. Popoola SI, Tsado Y, Ogunjinmi AA, Sanchez-Velazquez E, Peng Y, Rawat DB. Multi-stage deep learning for intrusion detection in industrial Internet of Things. IEEE Access. 2025;13(3):60532–55. doi:10.1109/access.2025.3557959.
27. Mendonça RV, Silva JC, Rosa RL, Saadi M, Rodriguez DZ, Farouk A. A lightweight intelligent intrusion detection system for industrial Internet of Things using deep learning algorithms. Expert Syst. 2022;39(5):e12917. doi:10.1111/exsy.12917.

28. Shtayat MM, Hasan MK, Sulaiman R, Islam S, Khan AUR. An explainable ensemble deep learning approach for intrusion detection in industrial Internet of Things. IEEE Access. 2023;11(2):115047–61. doi:10.1109/access.2023.3323573.
29. Nandanwar H, Katarya R. Deep learning enabled intrusion detection system for Industrial IOT environment. Expert Syst Appl. 2024;249(11):123808. doi:10.1016/j.eswa.2024.123808.
30. Kalinin M, Krundyshev V. Security intrusion detection using quantum machine learning techniques. J Comput Virol Hacking Tech. 2023;19(1):125–36. doi:10.1007/s11416-022-00435-0.
31. Shaji AA, Sadhwani S, Muthalagu R, Pawar PM, Mathew T. QML-IDS: quantum machine learning approaches for intrusion detection systems in IoT devices. In: Proceedings of the 2025 3rd International Conference on Computational Intelligence and Network Systems (CINS); 2025 Nov 25–26; Dubai, United Arab Emirates. p. 1–7.
32. Cirillo F, Esposito C, Taek Seo J. Benchmarking quantum machine learning methods for intrusion detection on noisy quantum computers. Quantum Mach Intell. 2026;8(1):27. doi:10.1007/s42484-026-00379-4.
33. Salek MS, Biswas PK, Pollard J, Hales J, Shen Z, Dixit V, et al. A novel hybrid quantum-classical framework for an in-vehicle controller area network intrusion detection. IEEE Access. 2023;11:96081–92. doi:10.36227/techrxiv.21907443.
34. Elsedimy E, Elhadidy H, Abohashish SM. A novel intrusion detection system based on a hybrid quantum support vector machine and improved Grey Wolf optimizer. Clust Comput. 2024;27(7):9917–35. doi:10.1007/s10586-024-04458-8.
35. Kuo SY, Shen JY, Liu CL, Chou YH. Hybrid quantum-inspired evolutionary neural networks for intrusion detection system. In: Proceedings of the 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC); 2024 Oct 6–10; Kuching, Malaysia. p. 2801–6.
36. Amara M, Mnasri S, Val T. QCNN-ID: a quantum-classical hybrid model for IoT intrusion detection. Procedia Comput Sci. 2025;270:2196–204.
37. Salvakkam DB, Saravanan V, Jain PK, Pamula R. Enhanced quantum-secure ensemble intrusion detection techniques for cloud based on deep learning. Cogn Comput. 2023;15(5):1593–612. doi:10.1007/s12559-023-10139-2.
38. Ishtiaq W, Zannat A, Shahariar Parvez AHM, Alamgir Hossain M, Hasan Kanchan M, Masud Tarek M. CST-AFNet: a dual attention-based deep learning framework for intrusion detection in IoT networks. Array. 2025;27:100501.
39. Alshehri MS, Saidani O, Alrayes FS, Abbasi SF, Ahmad J. A self-attention-based deep convolutional neural networks for IIoT networks intrusion detection. IEEE Access. 2024;12:45762–72. doi:10.1109/access.2024.3380816.
40. Hasan T, Hossain A, Ansari MQ, Syed TH. Enhanced intrusion detection in IIoT networks: a lightweight approach with autoencoder-based feature learning. arXiv:2501.15266. 2025.
41. Ullah S, Boulila W, Koubâa A, Ahmad J. MAGRU-IDS: a multi-head attention-based gated recurrent unit for intrusion detection in IIoT networks. IEEE Access. 2023;11:114590–601.