Explainable AI-Enabled Intrusion Detection in IoT Networks Using Optimized Quantum Reinforcement Learning

Surendra Kumar, *Member, IEEE*, Mridula Dwivedi, Mohit Kumar, Huaming Wu, *Senior Member, IEEE*, and Sukhpal Singh Gill

***Abstract*—Internet of Things (IoT) networks face increasing security vulnerabilities, necessitating the development of effective Intrusion Detection Systems (IDS). While Machine Learning (ML), Deep Learning (DL), and Reinforcement Learning (RL) have improved anomaly detection, existing approaches still suffer from high computational costs, limited adaptability to emerging attack patterns, and inefficiencies in handling high-dimensional IoT data, which restrict their real-time applicability. In this paper, we propose an IDS framework called QRL-SHO-XGBoost, which integrates Sea-horse Optimization (SHO) for efficient feature selection, Quantum Reinforcement Learning (QRL) for enhanced pattern recognition, and an XGBoost classifier for robust anomaly detection. Evaluations on benchmark datasets demonstrate superior performance, with testing accuracies of 99.20% on UNSW-NB15, 99.21% on CIC-DDoS2019, and 99.53% on AppLayer DDoS, while maintaining millisecond-level inference times. The quantum experiments analyze on 4–8 qubits shows consistently high fidelities ranging from 0.9815 to 0.9954. Although circuit depth increased from 12 to 24, the model maintained strong performance, with runtime increasing from 1.54 to 10.15 hours. Furthermore, model predictions are interpreted using Local Interpretable Model-Agnostic Explanations (LIME), enhancing transparency and interpretability. This research establishes a new benchmark for IoT IDS, offering a scalable, efficient, and explainable solution to protect critical infrastructures against threats in Cyber-Physical Systems (CPS).**

***Index Terms*—Sea-horse optimizer, local interpretable model-agnostic explanation, quantum reinforcement learning, XGBoost, intrusion detection system, Internet of Things.**

Received 1 March 2026; revised 26 June 2026; accepted 16 July 2026. Date of publication 22 July 2026; date of current version 13 November 2026. This work was supported in part by the Natural Science Foundation of Tianjin under Grant 25JCYBJC01540, in part by the National Key Research and Development Program of China under Grant 2025YFC3409900, and in part by the Emerging Frontiers Cultivation Program of Tianjin University Interdisciplinary Center. *(Corresponding author: Huaming Wu.)*

Surendra Kumar is with the Department of Computer Engineering and Applications, GLA University, Mathura 281406, India (e-mail: kumar.surendra1989@gmail.com).

Mridula Dwivedi is with the Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow 206025, India (e-mail: md.cs.bbau@gmail.com).

Mohit Kumar is with the Department of IT, Dr. B. R. Ambedkar National Institute of Technology, Jalandhar 144008, India (e-mail: kumarmohit@nitj.ac.in).

Huaming Wu is with the Center for Applied Mathematics and KL-AAGDM, Tianjin University, Tianjin 300072, China (e-mail: whming@tju.edu.cn).

Sukhpal Singh Gill is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS London, U.K. (e-mail: s.s.gill@qmul.ac.uk).

Digital Object Identifier 10.1109/TCE.2026.3715996

I. Introduction

THE Cyber-Physical Systems (CPS) integrate control systems, communication networks, sensors, and actuators into distributed architectures [1]. They are large-scale, heterogeneous, geographically distributed, and play a critical role in life-critical applications [2] including smart grids, drones, healthcare, and first responder awareness [3] where CPS operations require timing accuracy, predictable network traffic, and reliable wireless components [4]. Adversaries can use short- and long-term strategies [5] to plan distributed attacks without disrupting activities [6]. CPS security requires Intrusion Detection Systems (IDS) for intrusion detection. Machine Learning (ML)-based IDS improves detection accuracy and efficiency, while anomaly-based IDS models set a baseline and highlight deviations [7]. It is computationally intensive to monitor big and diversified network data [8]. Deep Learning (DL) enhances packet inspection, correlation finding, and zero-day detection [9], [10]. Reinforcement Learning-based IDS (RL-IDS) uses reinforcement learning algorithms to train an agent to detect intrusions by maximizing a reward signal [11]. Quantum computing (QC) enhances intrusion detection [12], where quantum networks offer unique opportunities and problems, so Quantum Machine Learning (QML) applies supervised, unsupervised, and reinforcement learning to complicated domains [13]. Quantum encoding and kernels enable QML to handle large datasets more efficiently than conventional ML [14], but state breakdown and decoding hinder performance monitoring [15]. Quantum Reinforcement Learning (QRL) models confront hardware noise, preprocessing complexity, and explainability issues for use with quantum circuits for decision-making [12]. Encoders, variational circuits, and measurement components complicate their architectures and confound interpretability [16]. Circuit entanglement obscures component contributions, reducing transparency [17]. Explainable Artificial Intelligence (XAI) improves trust and transparency to overcome these restrictions. To enable real-time IDS in healthcare and banking, XAI explains decision logic [18]. Reasoning and ML improve detection in explainable IDS setups. Evaluating feature importance with model-agnostic techniques like Local Interpretable Model-Agnostic Explanations (LIME) improves interpretability and transparency in classical and quantum-based IDS [19].

1558-4127 © 2026 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

Internet of Things (IoT) systems face major security challenges due to resource limitations and increasing exposure to DDoS attacks [20]. To improve intrusion detection in high-dimensional environments, several metaheuristic optimization algorithms, such as Particle Swarm Optimization (PSO), Grey Wolf Optimizer (GWO), Harris Hawks Optimization (HHO), Sea-Horse Optimization (SHO), have been explored for feature selection [21]. Previous studies have shown that SHO provides a good balance between exploration and exploitation and performs competitively compared to other optimization methods [22]. Based on this motivation, the proposed framework combines SHO with quantum-enhanced reinforcement learning and XAI to improve intrusion detection performance, reduce computational overhead, and provide transparent decision-making in IoT environments. The primary contributions of this work are as follows:

- This work integrates the bio-inspired SHO for feature selection and combines it with QRL for adaptive pattern recognition, alongside Extreme Gradient Boosting (XGBoost) for robust classification.
- The proposed model is evaluated on UNSW-NB15, CIC-DDoS2019, and AppLayer DDoS datasets, achieving accuracies of 99.20%, 99.21%, and 99.53%, respectively. Precision, recall, and F1-scores above 99.19% demonstrate efficient detection of normal and attack traffic with minimal false positives and false negatives.
- Proposed method with feature encoding and reinforcement learning high-fidelity quantum feature representations (0.9815–0.9954) across 4–8 qubits, improving adaptability and accuracy. XAI results demonstrating the interpretability of the proposed model.

The remainder of this article is structured as follows: Section II reviews key challenges and related work; Section III presents a novel IDS leveraging Q-learning with SHO (QRL-SHO); Section IV discusses experimental results alongside interpretability insights using XAI; and Section V concludes the paper while outlining future research perspectives.

II. Related Work

Quantum algorithms require data encoding to connect classical and quantum domains, enhance model performance, and increase classification accuracy and F1 scores [23]. Implementing these methods causes resource optimization, scalability, and information retention challenges [24]. In quantum technology, encoding methods must balance computational complexity and limited resources [25]. As data dimensions increase, quantum algorithms may lose their theoretical benefits due to encoding complexity [26]. Preserving important data properties during transformation is essential for model correctness and dependability [27].

There is a lack of studies comparing encoding strategies, creating a research gap. Encoding algorithms involve trade-offs, such as computational cost and scalability, despite being optimized for specific purposes [34]. These constraints complicate the adoption of encoding methods in quantum machine learning. Additionally, limited resources and noise in quantum hardware pose challenges for network intrusion detection [35].

TABLE I
Comparison Analysis of This Study With State-of-the-Arts

Ref	Quantum Encoding	Resource Efficiency	Multi-Threat Generaliz-ability	XAI	Low Time Complexity	Scalability (High-Dim)
[28]	✗	✗	✗	✗	✗	✗
[15]	✓	✗	✗	✗	✗	✗
[23]	✓	✗	✗	✗	✗	✓
[29]	✗	✗	✗	✗	✓	✗
[30]	✗	✗	✗	✗	✗	✗
[31]	✓	✗	✗	✗	✗	✓
[32]	✗	✗	✗	✗	✗	✗
[33]	✗	✗	✗	✗	✗	✗
This Study	✓	✓	✓	✓	✓	✓

To detect Distributed Denial of Service (DDoS) attacks, Dimensionality Reduction and Encoding Technique for Quantum Support Vector Classifier (DEQSVC), a quantum feature map, is applied to the CIC-DDoS2019 dataset. Qiskit and IBM's quantum computer simulations' 99.49% accuracy is impressive. Single data encoding limits alternate methodology research [29]. Quantum Support Vector Machines (QSVMs) and hybrid quantum-classical neural networks are used to detect DDoS threats using angle encoding on a dataset [30]. Security risk category needs validation across many cybersecurity challenges. Discusses theoretical and actual encoding and model performance [32]. Create hybrid quantum-classical models and variational circuits to optimise QML data encoding for hardware constraints. Encoding unique usage scenarios is easier with QML [31].

Traditional methods favor AI system openness and trustworthiness over explainability and interpretability, which aid higher-level aims. Investigate post-hoc interpretability methods such as Partial Dependence Plot (PDP), Accumulated Local Effects (ALE), Individual Conditional Expectation (ICE), Local Interpretable Model-agnostic Explanations (LIME), and SHapley Additive exPlanations (SHAP) [36] Closed box model interpretation used LIME and SHAP. Effects of disturbed samples and picture areas affect LIME predictions [37]. SHAP predicts updates to input features using game theory's Shapley values [38]. Table I shows the comparision of the propsoed study with exisitng works based on key parameters.

III. Proposed Methodology

A. Metaheuristic-Based Intrusion Detection

Traditional IDS systems have significant false alarm rates, computing costs, and inefficiencies in identifying sophisticated attacks, despite protecting IoT networks from cyber threats. The metaheuristic-driven SHO algorithm improves IDS performance. SHO balances exploration with exploitation, maximizing feature selection and intrusion detection accuracy, inspired by Hippocampus (Sea-horse) camouflage and hunting strategies. To apply SHO in a binary selection context we use a continuous relaxation $z \in \mathbb{R}^d$ with a sigmoid threshold to obtain $s_j = \text{Bernoulli}(\sigma(z_j))$ or simply $s_j = \mathbb{I}[\sigma(z_j) > \tau]$, where $\sigma(t) = (1+e^{-t})^{-1}$. A generic SHO positional update for individual i at iteration t can be written as:

$$v_i^{t+1} = \omega(t)\, v_i^t + \beta_1(t)\, R_1 \odot (p_i^t - z_i^t) + \beta_2(t)\, R_2 \odot (g^t - z_i^t) + \eta(t)\, J_i^t, \quad (1)$$

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

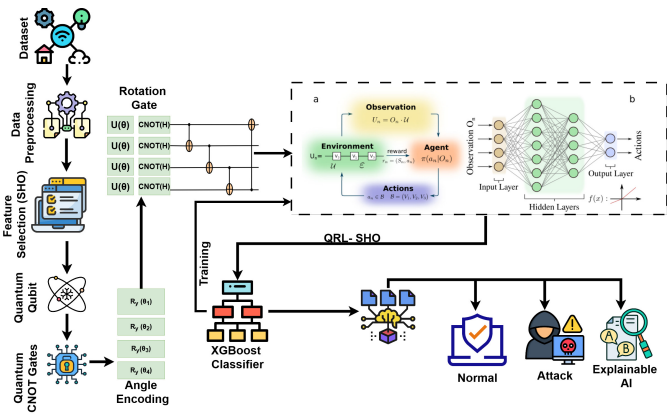


Fig. 1. Proposed quantum-based RL-SHO using XGBoost framework.

$$z_i^{t+1} = z_i^t + v_i^{t+1} \tag{2}$$

where $z_i^t \in \mathbb{R}^d$ is the continuous position of agent i, v_i^t its velocity-like term, p_i^t its personal-best position, g^t the global-best position, $R_1, R_2 \sim \mathcal{U}(0,1)^d$ are random vectors, $\odot$ denotes element-wise multiplication, and J_i^t is a stochastic jump vector (captures SHO's jump/adaptive camouflage). Coefficients $\omega, \beta_1, \beta_2, \eta$ may be time-varying to balance exploration/exploitation. After the update, convert to selection:

$$s_{i,j}^{t+1} = \mathbb{I}\{\sigma(z_{i,j}^{t+1}) > \tau\}, \quad j = 1, \ldots, d, \tag{3}$$

where $\tau \in (0,1)$ is a threshold (commonly $\tau = 0.5$).

Given population size N and maximum iterations T, for each agent i evaluate $\mathcal{F}(s_i)$ and update personal/global bests. The output $s^\star$ yields the final feature set $\mathcal{S}^\star = \{j : s_j^\star = 1\}$. SHO-based IDS dynamically detects essential features, reduces redundancy, and streamlines feature selection for classifiers like XGBoost. Adjustable search algorithms quickly converge to optimal security rules, improving attack detection and lowering false positives. The quantum-inspired pattern identification and anomaly detection of QRL-SHO integration improves IDS. To identify zero-day assaults, IDS SHO learns complicated attack patterns in real time. It analyzes attack behavior using SHO, XAI approaches like LIME, and strong ML models. Tests on real-world datasets show that SHO-based IDS reduces security threats and optimises IoT resource use as shown in Fig. 1.

B. Quantum Learning (QL)

IDS requires efficient feature transformation techniques to differentiate between normal and malicious network behavior. Quantum computing offers an innovative way to enhance feature representation and classification.

Let $x_\mathcal{S} \in \mathbb{R}^k$ be the k-dimensional reduced vector (after selection). We map classical features to quantum states via a feature map $U_\Psi(x_\mathcal{S}; \varphi)$ applied to the initial state $|0\rangle^{\otimes\zeta}$ and $\zeta = 8$:

$$|\psi(x_\mathcal{S}; \varphi)\rangle = U_\Psi(x_\mathcal{S}; \varphi)\,|0\rangle^{\otimes\zeta}, \tag{4}$$

We extract classical descriptors from the quantum state using expectation values in the Pauli-Z basis:

$$z_q(x_\mathcal{S}) = \langle\psi(x_\mathcal{S}; \varphi)|Z_q|\psi(x_\mathcal{S}; \varphi)\rangle, \quad q = 1, \ldots, \zeta, \tag{5}$$

Algorithm 1 QRL-Driven IDS

Input: Dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^M$
Output: Optimized intrusion detection model μ_{final}
procedure LOADDATASET(ϕ)
  $\Delta \leftarrow$ read_csv(ϕ)
  return Δ
end procedure
procedure PREPROCESSDATA(Δ)
  for $\alpha \in$ CategoricalColumns(Δ) **do**
    if $\alpha \neq$ Label **then**
      $\Delta[\alpha] \leftarrow$ LabelEncoder($\Delta[\alpha]$)
    end if
  end for
  $\Gamma \leftarrow$ NumericColumns(Δ) $\setminus$ {Label}
  $\Delta[\Gamma] \leftarrow$ MinMaxScaler($\Delta[\Gamma]$)
  Δ[Label] $\leftarrow$ LabelEncoder(Δ[Label])
  return Δ, Γ
end procedure
Quantum Backend: 8-qubit simulator
procedure QUANTFEATUREMAP($\mathbf{x}_S$)
  for $k = 1$ to 8 **do**
    Normalize feature: $\theta_k = \pi \cdot x_k$
    Encode feature using angle encoding: $|\psi_k\rangle = R_Y(\theta_k)|0\rangle$
  end for
  Apply first entanglement layer:
  $CNOT(1,2)$, $CNOT(2,3)$, $CNOT(3,4)$,
  $CNOT(4,5)$, $CNOT(5,6)$, $CNOT(6,7)$, $CNOT(7,8)$
  Apply ring entanglement:
  $CNOT(8,1)$
  Apply variational rotations:
  for $k = 1$ to 8 **do**
    $R_X(\phi_k)$
    $R_Z(\omega_k)$
  end for
  Measure expectation values: $z_k = \langle\psi|Z_k|\psi\rangle$
  return $\mathbf{z} = [z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8]$
end procedure
procedure APPLYQUANTUMTRANSFORM($\mathcal{X}_S$)
  $\mathcal{X}^q \leftarrow$ {QuantFeatureMap(x_i) | $x_i \in \mathcal{X}_S$}
  return $\mathcal{X}^q$
end procedure
procedure QRL-OPTIMIZE($\mathcal{X}^q_{\text{train}}, \mathcal{Y}_{\text{train}}$)
  Initialize SHO population size N_p
  Initialize Q-table $Q(s,a)$
  for $t = 1$ to $T_{\max}$ **do**
    for each horse i **do**
      Train XGBoost with parameters $\theta_i^{(t)}$
      Evaluate validation F1-score $\text{F1}_{\text{val}}^{(t)}$
      Define state: $s_t = \{\text{F1}_{\text{val}}^{(t)}, |\mathcal{S}|, \theta_i^{(t)}\}$
      Select action a_t using ϵ-greedy policy
      Update parameters: $\theta_i^{(t+1)} = \theta_i^{(t)} + a_t$
      Compute reward: $r_t = \text{F1}_{\text{val}}^{(t)} - \lambda\frac{|\mathcal{S}|}{d}$
      Update Q-table:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\big[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)\big]$$

    end for
    if convergence criterion satisfied **then**
      break
    end if
  end for
  return θ^*
end procedure

These extracted expectation values or similarity measures for the downstream classifier.

1) Quantum Feature Encoding Using Rotation Gates: IoT network traffic features (such as packet size, connection

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

duration, or anomaly scores) are fed into machine learning models for classification. In quantum computing, these features are encoded into qubit states using quantum rotation gates RX:

$$R_X(\theta) = e^{-i\frac{\theta}{2}X} = \begin{bmatrix} \cos 2\theta & -i\sin 2\theta \\ -i\sin 2\theta & \cos 2\theta \end{bmatrix}. \tag{6}$$

This rotates the qubit state around the X-axis, transforming the probability amplitudes.

$$R_Y(\theta) = e^{-i\frac{\theta}{2}Y} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{bmatrix}. \tag{7}$$

This controls the balance between qubit states, helping in encoding probabilities.

$$R_Z(\theta) = e^{-i\frac{\theta}{2}Z} = \begin{bmatrix} e^{-i2\theta} & 0 \\ 0 & e^{i2\theta} \end{bmatrix}. \tag{8}$$

This phase shift impacts interference between qubits. Each feature x_i from an IDS dataset (e.g., packet size, source IP entropy) is mapped to an angle θ_i, transforming an initial state $|\psi\rangle$ into a quantum representation:

$$|\psi\rangle = R_X(x_0)R_Y(x_1)R_Z(x_2)|0\rangle. \tag{9}$$

This encoding captures non-linear dependencies in the data, improving IDS performance by enabling high-dimensional quantum feature transformations.

2) CNOT Gates: Entanglement is crucial for IDS applications as it enables qubits to share information, capturing dependencies between network features. The Controlled-NOT (CNOT) gate introduces entanglement between qubits. Its matrix form is:

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{10}$$

Packet size, connection time, source and destination communication patterns, protocol statistics, flow entropy, payload characteristics, and anomaly scores are interdependent for intrusion detection in IoT networks. A ring-topology CNOT design provides an entanglement layer to capture these correlations in quantum feature space. The quantum register propagates local feature correlations by entangling neighboring qubits using CNOT operations between pairings (q_1, q_2), (q_2, q_3), …, and (q_7, q_8). The circuit models global interdependence among distributed network traffic attributes after a further CNOT operation between qubits q_8 and q_1 closes the entanglement loop. The following quantum state comes from this ring-entangled structure's depiction of complicated, nonlinear intrusion-related feature relationships:

$$|\Psi(\mathbf{x}, \boldsymbol{\phi})\rangle = \left[\prod_{l=1}^{L} U_{\text{var}}^{(l)}(\boldsymbol{\phi}) U_{\text{ent}}^{(l)}\right] U_{\text{enc}}(\mathbf{x})|0\rangle^{\otimes 8}, \tag{11}$$

$$U_{\text{enc}}(\mathbf{x}) = \bigotimes_{i=1}^{8} R_Y(\theta_i) \tag{12}$$

$$U_{\text{ent}}^{(l)} = \prod_{i=1}^{8} CNOT\ (i,\ (i \bmod 8) + 1) \tag{13}$$

$$U_{\text{var}}^{(l)} = \prod_{i=1}^{8} R_Z(\phi_{i,l}^{(3)}) R_Y(\phi_{i,l}^{(2)}) R_X(\phi_{i,l}^{(1)}) \tag{14}$$

Through entanglement, SHO-IDS can detect subtle patterns in IoT network traffic anomalies that classical feature transformation methods may overlook.

3) Pauli-Z Basis: Once feature encoding and entanglement are complete, measurement is performed to extract classical information for intrusion classification. In quantum systems, measurements collapse the quantum state into a classical outcome. The Pauli-Z measurement operator is defined as:

$$\sigma_Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{15}$$

where the expectation value of Pauli-Z for a quantum state $|\psi\rangle$ is:

$$\langle \sigma_Z \rangle = \langle \psi | \sigma_Z | \psi \rangle. \tag{16}$$

Each expectation value reflects an encoded network feature, aiding in classifying traffic as normal or malicious. By leveraging quantum feature transformation, IDS can enhance attack detection accuracy, capturing relationships that classical methods might miss. The selected features are mapped one-to-one to quantum qubits using angle encoding. Each feature $x_i \in \mathbf{x}_S$ is normalized to $[0, \pi]$ and encoded as a rotation angle:

$$\theta_i = \pi \cdot \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}, \quad i = 1, \ldots, 8. \tag{17}$$

Each qubit q_i encodes exactly one classical feature via a parameterized rotation gate:

$$|\Psi\rangle = \left(\prod_{i=1}^{7} CNOT(i, i+1)\right) CNOT(8, 1) \times \left(\bigotimes_{i=1}^{8} R_Y(\theta_i)\right) |0\rangle^{\otimes 8} \tag{18}$$

Thus, the 8 qubit quantum state represents a compressed embedding of the original high-dimensional traffic feature vector. The reported quantum fidelity measures the similarity between the ideal (noise-free) quantum feature-embedding state ρ_{ideal} and the noisy, hardware-realistic state ρ_{noisy}:

$$F(\rho_{\text{ideal}}, \rho_{\text{noisy}}) = \left(\text{Tr}\sqrt{\sqrt{\rho_{\text{ideal}}}\rho_{\text{noisy}}\sqrt{\rho_{\text{ideal}}}}\right)^2. \tag{19}$$

Qiskit Aer Simulator is used to run all quantum experiments on an 8-qubit variational quantum circuit with three encoding-entanglement layers. For statistically consistent expectation-value estimate, each circuit execution has 8192 measurement shots. A depolarized noise model, measurement-error mitigation, and Zero-Noise Extrapolation (ZNE) assessed fidelity in noise-aware simulations. After transpilation and circuit optimization, the 24-depth quantum circuit included 72 single-qubit rotation gates and 24 two-qubit entangling gates, providing 96 quantum operations. About 1500 reinforcement learning episodes are run, with frequent quantum circuit evaluations for policy updates, reward estimates, and fidelity calculation. The 8-qubit state space and deeper entanglement structure increased computational complexity, resulting in a

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

9.85-hour simulator runtime. Despite an 8-qubit circuit execution taking only milliseconds, repetitive circuit sampling, reinforcement learning optimization, and error-mitigation processes increased training time. Quantum fidelity is the average state fidelity between the ideal noise-free quantum embedding and the noise-affected embedding after measurement-error mitigation and ZNE across all reinforcement learning events.

Hyperparameter tuning and adaptive control are defined as a Markov Decision Process (MDP) $(\mathcal{S}_{\mathrm{RL}}, \mathcal{A}, P, R, \gamma)$, where $\mathcal{S}_{\mathrm{RL}}$ is the RL state space (e.g., current hyperparameter vector, recent validation loss/f1). $\mathcal{A}$ is the action space (discrete or continuous adjustments to hyperparameters θ or SHO control params). $R(s,a)$ is a reward reflecting classifier quality and sparsity, e.g. $R = \mathrm{F1}_{\mathrm{val}} - \lambda \cdot \frac{\|s\|_1}{d}$. $\gamma \in [0,1)$ is the discount factor. A Q-learning style update for a tabular or discretized Q-function is:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\big(r_t + \gamma \max_{a'} Q(s_{t+1},a') - Q(s_t,a_t)\big), \tag{20}$$

where α is the learning rate. For continuous actions or parameterized policies, use actor-critic or policy gradient variants where policy $\pi_\phi(a|s)$ is updated to maximize expected return:

$$\nabla_\phi J(\phi) = \mathbb{E}_{\tau\sim\pi_\phi}\left[\sum_{t=0}^{T} \nabla_\phi \log \pi_\phi(a_t|s_t)\, G_t\right], \tag{21}$$

with return $G_t = \sum_{k=t}^{\infty} \gamma^{k-t} r_k$. If the policy or value estimator is parameterized by a PQC, denote $V_\varphi(s) \approx \langle M \rangle_{\varphi,s}$ where M is an observable; training then adjusts φ to minimize temporal-difference (TD) loss

$$\mathcal{L}_{\mathrm{TD}}(\varphi) = \mathbb{E}_{(s,a,r,s')}\big[\big(r + \gamma V_\varphi(s') - V_\varphi(s)\big)^2\big]. \tag{22}$$

Reinforcement learning uses discretized optimization states from validation performance, selected-feature cardinality, and optimization parameters. In each iteration, the agent observes state s_t, chooses action a_t using a ϵ-greedy strategy, obtains reward r_t, and updates the Q-table using temporal-difference learning. Rewards are the mean of independent runs with varied random initializations. While the baseline RL achieved a cumulative reward of 236.1 $\pm$ 5.1, the proposed QRL framework obtained 245.8 $\pm$ 4.3. A paired two-sided t-test confirmed a substantial improvement (p‘!0.05). Q-learning agent is set to a 0.10 learning rate (α) and 0.95 discount factor (γ). An ϵ-greedy exploration approach is established, starting at 1.0 and decreasing by 0.995 to a minimum of 0.01. An action space with five discrete optimization actions trained the agent for 1500 episodes.

$$s_t = \mathrm{F1}_{\mathrm{val}}^{(t)}, |\mathcal{S}|, \theta^{(t)}, \tag{23}$$

C. XGBoost Classifier-Based QRL

Let the ensemble consist of T regression trees $f_t(x)$; XGBoost minimizes the regularized objective:

$$\mathcal{L}_{\mathrm{XGB}} = \sum_{i=1}^{n} \ell\big(y_i, \hat{y}_i^{(T)}\big) + \sum_{t=1}^{T} \Omega(f_t), \tag{24}$$

Algorithm 2 Quantum-Enhanced Feature Selection and RL-SHO (QRL-SHO) Optimized XGBoost

Input: Quantum-transformed feature matrix $\mathcal{X}^q$, labels $\mathcal{Y}$
Output: Optimized XGBoost model μ_{final}
procedure FEATURES($\mathcal{X}, \mathcal{Y}, k$)
  Train preliminary XGBoost model μ using $(\mathcal{X}, \mathcal{Y})$

$$\phi_j = \frac{1}{m}\sum_{i=1}^{m} \mathbb{I}(\text{Feature}_j \text{ is used in tree } i)$$

  $\mathcal{X}_{\mathrm{selected}} \leftarrow \mathcal{X}[:, \text{top-}k(\phi)]$
  return $\mathcal{X}_{\mathrm{selected}}$
end procedure
$\mathcal{X}_{\mathrm{train}}^{\mathrm{selected}} \leftarrow \mathrm{Features}(\mathcal{X}_{\mathrm{train}}^{q}, \mathcal{Y}_{\mathrm{train}}, k = 8)$
procedure RL-SHO($\mathcal{X}, \mathcal{Y}, N, T_{\max}$)
  Initialize N SHO horses (solutions) $\theta_i^{(0)} \sim \mathcal{U}(\Theta)$
  Initialize Q-table $Q(s,a)$ arbitrarily
  Set convergence threshold ε and patience window K
  for $t = 1$ to $T_{\max}$ **do**
  for each horse d_i **do**
    Train XGBoost with parameters $\theta_i^{(t)}$
    Evaluate validation F1-score $\mathrm{F1}_{\mathrm{val}}^{(t)}$

$$s_t = \{\mathrm{F1}_{\mathrm{val}}^{(t)}, |\mathcal{S}^{(t)}|, \theta_i^{(t)}\}$$

    Select action $a_t \in \mathcal{A}$ using ϵ-greedy policy $\theta_i^{(t+1)} = \theta_i^{(t)} + \alpha a_t$

$$r_t = \mathrm{F1}_{\mathrm{val}}^{(t)} - \lambda\frac{|\mathcal{S}^{(t)}|}{d}$$

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\big[r_t + \gamma \max_{a'} Q(s_{t+1},a') - Q(s_t,a_t)\big]$$

  end for
  if $|\mathrm{F1}_{\mathrm{val}}^{(t)} - \mathrm{F1}_{\mathrm{val}}^{(t-1)}| < \varepsilon$ for K consecutive iterations **then**
    break
  end if
  end for
  return Best hyperparameters: $\theta^* = \arg\max_\theta \mathrm{F1}_{\mathrm{val}}(\theta)$
end procedure
$\theta^* \leftarrow \text{RL-SHO}(\mathcal{X}_{\mathrm{train}}^{\mathrm{selected}}, \mathcal{Y}_{\mathrm{train}}, N, T_{\max})$
Train final model: $\mu_{\mathrm{final}} \leftarrow \mathrm{XGBoost}(\mathcal{X}_{\mathrm{train}}^{\mathrm{selected}}, \mathcal{Y}_{\mathrm{train}}, \theta^*)$

where $\hat{y}_i^{(T)} = \sum_{t=1}^{T} f_t(x_i)$, ℓ is the classification loss (logistic loss), and the tree regularizer is

$$\Omega(f) = \gamma T_f + \tfrac{1}{2}\lambda \sum_{j=1}^{T_f} w_j^2, \tag{25}$$

where T_f departs and w_j weights. In additive training, ideal leaf weights and gain per split are calculated using the second-order Taylor expansion of ℓ. Quantum entanglement correlations reduce input data redundancy in the QRL-XGBoost framework's higher-dimensional Hilbert space, enhancing intrusion detection. XGBoost classifies carefully selected features by fixing mistakes. Fast convergence without overfitting with parallelization and quantum-inspired hyperparameter adjustment. Each input $\boldsymbol{\xi} \in \mathbb{R}^v$ is mapped into a

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

quantum Hilbert space $\mathcal{H}$ using a quantum encoding function $\Psi : \mathbb{R}^{\nu} \rightarrow \mathbb{C}^{2^{\zeta}}$:

$$\Psi(\boldsymbol{\xi}) = \mathbf{U}_{\Psi}|0\rangle^{\otimes\zeta}, \tag{26}$$

where $\mathbf{U}_{\Psi}$ is a unitary transformation applied to $|0\rangle^{\otimes\zeta}$ and $\zeta = 8$ represents the number of quantum qubits. The transformed quantum state is:

$$|\psi_{\xi}\rangle = \sum_{\alpha=1}^{\kappa} \lambda_{\alpha}|\xi_{\alpha}\rangle, \quad \lambda_{\alpha} = \tau_{\alpha}e^{i\varphi_{\alpha}}, \tag{27}$$

where λ_{α} represents complex probability amplitudes, and φ_{α} is the quantum phase factor. The feature selection score for each feature θ_{ω} is computed as:

$$\phi_{\omega} = \frac{1}{\beta}\sum_{\tau=1}^{\beta} \mathbb{I}(\theta_{\omega} \text{ appears in } \tau), \tag{28}$$

where β is the total number of decision trees used for feature evaluation.

Algorithm 3 XAI Based SHO-IDS

Input: Trained model μ_{final}, test dataset $\mathcal{X}_{\text{test}}^{\text{selected}}$, number of perturbations M
Output: Local explanations $\mathcal{E}$
Initialize LIME explainer:
explainer←LimeTabularExplainer($\mathcal{X}_{\text{train}}^{\text{selected}}$, feature_names)
for each instance $\chi \in \mathcal{X}_{\text{test}}^{\text{selected}}$ **do**
  $\mathcal{X}_{\chi}^{M} = \{\tilde{\chi}_i \mid \tilde{\chi}_i = \chi + \delta_i, \quad \delta_i \sim \mathcal{N}(0, \sigma^2)\}_{i=1}^{M}$
  $\mathcal{Y}_{\chi}^{M} = \{\mu_{\text{final}}(\tilde{\chi}_i) \mid \tilde{\chi}_i \in \mathcal{X}_{\chi}^{M}\}$
  $w_i = \exp\left(-\frac{\|\tilde{\chi}_i - \chi\|_2^2}{\sigma^2}\right), \quad \forall i \in \{1, \ldots, M\}$
  $\hat{g}_{\chi}(\mathbf{x}) = \arg\min_{g \in G} \sum_{i=1}^{M} w_i \cdot \mathcal{L}(g(\tilde{\chi}_i), \mathcal{Y}_{\chi}^{M})$
  $\epsilon_{\chi} = \nabla_{\mathbf{x}}\hat{g}_{\chi}(\mathbf{x})\Big|_{\mathbf{x}=\chi}$
  $\mathcal{E} \leftarrow \mathcal{E} \cup \{\epsilon_{\chi}\}$
end for
return $\mathcal{E}$

D. Quantum Reinforcement SHO With XGBoost

Quantum feature selection, reinforcement learning (RL), SHO, and XGBoost provide a robust and adaptive pipeline. After quantum transformation and feature ranking, the top-k features are trained. RL-SHO module optimizes horse configurations using ϵ-greedy Q-learning strategy, SHO-inspired position updates, and XGBoost loss function rewards. Fig. 2 shows, a converged XGBoost model that balances accuracy, efficiency, and robustness for complex, high-dimensional tasks through continuous iterations. The proposed method increases generalization and scalability for real-world classification efficiency and accuracy. Quantum update using RL:

$$Q(\varsigma, \delta) \leftarrow Q(\varsigma, \delta) + \alpha[\rho + \gamma \max_{\delta'} Q(\varsigma', \delta') - Q(\varsigma, \delta)], \tag{29}$$

where ρ is the reward, γ is the discount factor, and Q is the quantum-influenced Q-value. Once the optimized model is trained, the final prediction function $\hat{\gamma}$ is given by:

$$\hat{\gamma} = \sum_{\tau=1}^{\lambda} f_{\tau}(\theta) + \sum_{\omega=1}^{\delta} \mathbf{K}_{\Psi}(\theta_{\omega}, \theta), \tag{30}$$

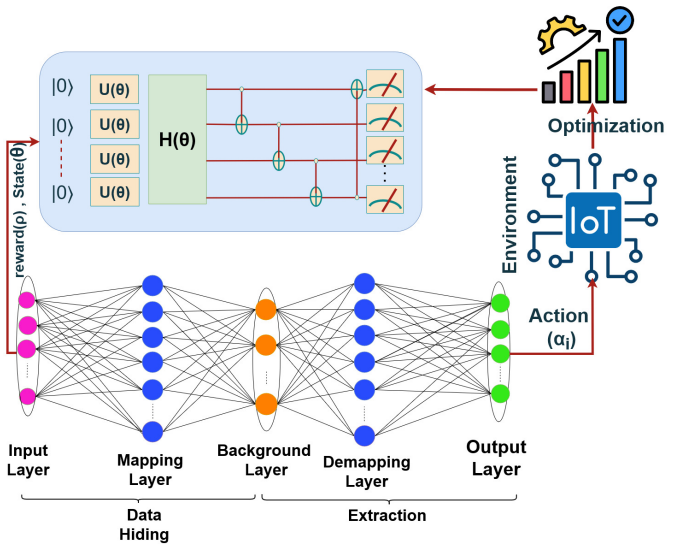


Fig. 2. Proposed quantum reinforcement learning integrates with SHO.

where the first summation represents the ensemble of boosted decision trees, and the second term represents the quantum-enhanced prediction.

E. Quantum Feature Transformation Impact

If $\mathcal{X}$ represents input features and $\mathcal{Y}$ represents labels, the quantum feature encoding applies Rotation Gates (R_x, R_y, R_z) and CNOT Gates to map classical data into quantum states:

$$R_X(\theta) = e^{-i\frac{\theta}{2}X}, \; R_Y(\theta) = e^{-i\frac{\theta}{2}Y}, \; R_Z(\theta) = e^{-i\frac{\theta}{2}Z}, \tag{31}$$

where CNOT gates establish entanglement among qubits, improving representation capability. After applying quantum transformations, the dataset $\mathcal{X}^{q}$ is extracted, containing both normal and attack traffic data. Feature engineering removes redundant values, normalizes data, and improves model efficiency. RL-SHO selects informative features by optimizing a feature subset iteratively:

$$\theta_i^{(t+1)} = \theta_i^{(t)} + \alpha a_i, \tag{32}$$

where actions a_i are guided by an adaptive Q-learning approach. Quantum embeddings encode features in a high-dimensional Hilbert space. Measurements in the Pauli-Z basis enhance separability between normal and attack instances. Hadamard and CNOT gates create a strongly entangled quantum layer. The transformation is modeled as:

$$U(\theta) = \prod_{i=1}^{L} R_x(\theta_i)R_y(\theta_i)R_z(\theta_i)CZ_{(i,i+1)}, \tag{33}$$

where L denotes the number of layers. The XGBoost model μ is trained on quantum-transformed data and optimized using RL-SHO. The classifier iteratively refines hyperparameters θ^{*} using:

$$\theta^{*} = \arg\max_{\theta} \mathbb{E}[Q(s, a; \theta)], \tag{34}$$

where $Q(s, a)$ is the Q-value function. Intrusion detection relies on fidelity computation between observed states ρ and trained normal states σ:

$$F = \text{Tr}(\sqrt{\rho}\sigma\sqrt{\rho})^2. \tag{35}$$

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

Proposed model enhance detection rules to respond to new threats and ensure intrusion detection. For a controlled comparison, we replaced the 8-qubit quantum feature map with a classical radial basis function (RBF) kernel while maintaining the reinforcement learning framework, training budget, and evaluation protocol. The proposed model outperforms the classical baseline with a final performance of 245.8, faster convergence (1.210^5 vs. 1.610^5 rounds), and decreased variance among seeds. This suggests that quantum feature mapping adds representational diversity rather than tuning hyperparameters. The enhancement in the hybrid learning framework is functional, not quantum.

IV. Results and Discussion

Qiskit Aer with the *aer_simulator* backend is used to simulate quantum experiments, as current Noisy Intermediate-Scale Quantum (NISQ) hardware does not support scalable deployment for the considered circuit depth [12]. The simulation included a depolarizing noise model and measurement error mitigation. Python libraries (NumPy, Pandas, Scikit-learn, Qiskit, Matplotlib, Seaborn) with shallow variational circuits ($L = 3$; RX–RY–RZ, sparse CNOT) are transpiled in Qiskit for depth/gate analysis in Google Colab experiments. After simulating NISQ noise (decoherence, gate/readout faults), mthree and Zero-Noise Extrapolation (mitiq) enhanced fidelity by 1.5–3%, stabilizing Pauli-Z expectations and entanglement correlations. Earlier IDS results show 97.5% (UNSW-NB15) [28], 98.2% (CIC-DDoS2019) [15], 98.7% (1.2 GB overhead) [23], and 98.9% (75% CPU) [33] SHO-QRL-XGBoost led with 99.20-99.53% accuracy, 1.2 ms inference, 67.3% CPU, 895 MB RAM, and ideal for real-time IoT IDS.

A. Dataset Description

UNSW-NB15[1] contains 2,540,044 network flow records, 49 features, 2,218,761 normal instances, 321,283 attack instances, and nine attack categories, including fuzzers, exploits, reconnaissance, and worms, making it suitable for general intrusion detection research. The Canadian Institute for Cybersecurity's CIC-DDoS2019 dataset,[2] with over 12,452,505 records and 88 extracted traffic features, includes statistical and protocol-based attributes. The dataset is significantly dispersed with 3,113,464 benign flows and 9,339,041 DDoS attack flows, including Synchronize (SYN), User Datagram Protocol (UDP), and HTTP floods. It accurately depicts high-volume cyber risks and provides traffic pattern information. The App-Layer DDoS dataset[3] has 3,085,728 records with 21 features, including 1,038,932 normal samples and 2,046,796 attack samples, focusing on Layer-7 HTTP-based DDoS attacks. Hash-based identification from the 5-tuple (source IP, destination IP, source port, destination port, and protocol) deleted redundant network flows to provide reliable evaluation and prevent data leakage.

[1]https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2]https://www.unb.ca/cic/datasets/ddos-2019.html
[3]https://www.kaggle.com/datasets/wardac/applicationlayer-ddos-dataset

TABLE II
Performance Comparison of Various Methods

Dataset	Phase	Class	Accuracy	Precision	Recall	F1-Score
UNSW-NB15	Training	Normal	98.76%	98.18%	99.64%	98.50%
		Anomaly	99.01%	98.14%	99.48%	98.45%
		Average	98.89%	98.16%	99.56%	98.46%
	Testing	Normal	99.20%	99.80%	98.60%	99.20%
		Anomaly	99.72%	98.76%	99.59%	99.30%
		Average	99.46%	99.28%	99.08%	99.25%
CIC-DDoS2019	Training	Normal	99.16%	99.18%	99.24%	99.09%
		Anomaly	99.10%	99.04%	99.38%	99.03%
		Average	99.13%	99.11%	99.31%	99.06%
	Testing	Normal	99.30%	99.10%	99.16%	99.18%
		Anomaly	99.12%	98.88%	99.56%	99.20%
		Average	99.21%	99.02%	99.36%	99.19%
App-DDoS	Training	Normal	98.96%	99.18%	99.04%	99.01%
		Anomaly	99.02%	99.14%	99.06%	99.02%
		Average	98.99%	99.16%	99.05%	99.04%
	Testing	Normal	99.46%	99.28%	99.26%	99.18%
		Anomaly	99.60%	99.34%	99.56%	99.19%
		Average	99.53%	99.31%	99.41%	99.19%

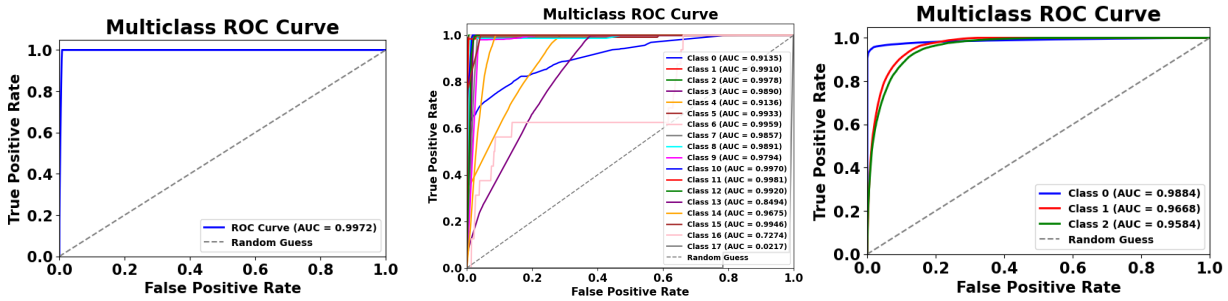
B. Dataset Preprocessing and Feature Engineering

UNSW-NB15, CIC-DDoS2019, and Application-Layer DDoS datasets are preprocessed before model training. The median statistics imputed missing values, Min-Max scaling normalized numerical features to $[0, 1]$, and one-hot encoding altered categorical characteristics. Class weighted learning in the XGBoost classifier optimizes underrepresented classes to reduce class imbalance. This technique increases minority-class misclassification penalties while retaining data distribution. In addition, stratified data partitioning is used to avoid overlap between training (70%), validation (15%), and test (15%) class proportions. Pre-processing and feature optimization using the SHO algorithm identified informative features and eliminated redundant attributes. Principal Component Analysis (PCA) and autoencoders are used to reduce latent representations from the feature subset, minimizing noise and computational complexity. Through variational quantum circuits, these representations are transferred onto a quantum feature space to extract higher-order feature correlations. Quantum encoding requires dimensionality reduction due to qubit availability and circuit depth constraints on existing quantum computing platforms.

C. Performance Analysis

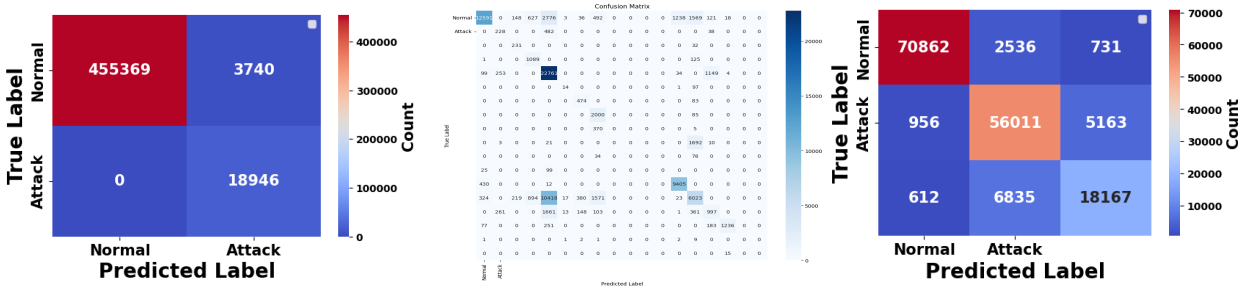
The framework is evaluated using accuracy, precision, recall, and F1-score. Classification performance is measured by accuracy. Precision is used to measure false-positive alarm reduction. Recall examines attack detection. Precision and recall are balanced in F1-score. All measures help assess the SHO-QRL-XGBoost intrusion detection framework's efficacy and reliability. Policy entropy, defined as: $H(\pi) = -\sum_{a\in\mathcal{A}} \pi(a|s) \log \pi(a|s)$, where $\pi(a|s)$ represents the probability of select action a in state s. Policies converge with lower entropy. Regret measures best-practice versus learnt policy performance. $\text{Regret}(T) = \sum_{t=1}^{T} \left(r_t^* - r_t\right)$, where r_t^* represents the greatest achievable reward (at iteration t) and r_t represents the Cumulative reward, which is the mean of all iterations' rewards. $(\%)CumulativeReward = \frac{\sum_{t=1}^{T} r_t}{\sum_{t=1}^{T} r_t^*} \times 100$.This measure shows how well the RL policy approaches optimal performance. Table III shows averages of RL measures across five separate runs to allow for learning stochasticity.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.



(a) UNSW Dataset ROC (b) CIC-DDoS Dataset ROC (c) App Layer Dataset ROC

Fig. 3. ROC analysis on various datasets.



(a) UNSW Confusion matrix (b) CIC-DDoS Confusion matrix (c) App Layer Confusion matrix

Fig. 4. Confusion matrix analysis on various datasets.

TABLE III

REINFORCEMENT LEARNING EVALUATION METRICS ACROSS DATASETS

Metric Category	Metric Name	UNSW-NB15	CIC-DDoS2019	App-DDoS
Reward-Based	Cumulative Reward (%)	98.5	98.3	98.8
	Average Reward	0.97	0.98	0.98
	Discounted Return	0.85	0.93	0.91
Policy Performance	Success Rate (%)	99.1	99.5	99.9
	Policy Entropy	0.14	0.09	0.16
	Regret	0.05	0.03	0.07
Efficiency	Episodes to Convergence	1200	950	1350
	Sample Efficiency (%)	97	98	95
Generalization & Robustness	Generalization Score (%)	98.2	98.6	99.1
	Transfer Learning Efficiency (%)	93	97	98
Stability & Convergence	Training Stability (%)	99.56%	99.2%	99.50%
	Loss Convergence Rate	0.07	0.06	0.08
Exploration vs. Exploitation	Exploration Rate	0.21	0.19	0.24
	Action Distribution	60/40 (Attack/Normal)	64/36 (Attack/Normal)	57/43 (Attack/Normal)

Compare baselines to see each component's contribution. SHO-XGBoost enhanced accuracy to 98.74% and F1-score to 98.68%. RL-XGBoost improved accuracy to 98.91% and F1-score (98.85%). The suggested SHO-QRL-XGBoost system outperformed others with 99.21% accuracy and 99.19% F1-score, demonstrating the overlap between feature optimization and quantum-enhanced reinforcement learning. This adaptation optimizes quantum-assisted model precision and recall for complex attack detection. Fig. 3 displays the performance evaluation ROC Curve, while Fig. 4 displays the comparison confusion matrix analysis.

The experimental results are evaluated against baseline methods reported in prior studies [15], [23], [28], [33]. The classical ML approach in [28] achieves 97.5% accuracy on the UNSW-NB15 dataset, while the quantum-enhanced IDS in [15] reaches 98.2% on CIC-DDoS2019 using angle encoding; however, both methods exhibit relatively high inference latency (5.8 ms and 4.2 ms, respectively) and limited interpretability. An encoding-focused method reports 98.7% accuracy but incurs substantial memory overhead (1.2 GB),

TABLE IV

COMPUTATIONAL LATENCY ANALYSIS OF SHO-IDS

Dataset	Preprocessing Time (ms)	Feature Selection Time (ms)	Training Time (s)	Inference Time (ms)
UNSW-NB15	12.4	8.7	34.2	1.2
CIC-DDoS2019	15.8	10.2	41.7	1.6
App-DDoS	10.9	6.5	29.3	1.0

which restricts its applicability in IoT environments [23]. Similarly, a variational circuit model achieves 98.9% accuracy but faces scalability challenges, requiring high CPU utilization (75%) [33]. In contrast, the proposed SHO-QRL-XGBoost framework achieves superior accuracies ranging from 99.20% to 99.53% across the UNSW-NB15, CIC-DDoS2019, and AppLayer DDoS datasets, while maintaining low inference latency (1.2 ms), moderate resource usage (67.3% CPU and 895 MB RAM), and improved interpretability through LIME. Circuit compilation, calibration, and classical quantum interface delay are only required during offline training and amortized across learning iterations. Quantum state tomography is not needed. Because an optimized XGBoost classifier with millisecond inference delay detects runtime, the proposed system facilitates real-time IoT deployment. Offline quantum feature encoding and reinforcement learning optimization during training do not increase operational inference overhead.

The proposed QRL-SHO-XGBoost framework performs quantum processing exclusively during the offline training stage, the runtime energy consumption is governed by classical inference. The energy consumed per inference is modeled as

$$\mathcal{E}_{\text{inf}} = \sum_{i \in \{\text{CPU,GPU}\}} (\rho_i \cdot \pi_i \cdot \tau_{\text{inf}}), \tag{36}$$

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

TABLE V
RESOURCE UTILIZATION AND ENERGY CONSUMPTION ANALYSIS (INFERENCE PHASE)

Dataset	CPU Usage (%)	RAM Consumption (MB)	GPU Utilization (%)	Energy per Inference (mJ)
UNSW-NB15	67.3	895	45.2	1.18
CIC-DDoS2019	73.9	1042	49.6	1.45
App-DDoS	58.7	780	40.3	0.96

TABLE VI
QUANTUM CIRCUIT SCALABILITY ANALYSIS

Qubits	Latent	Depth	1-Q Gates	2-Q Gates	Fidelity	Runtime (h)
4	2	12	36	6	0.9954	1.54
4	3	12	36	6	0.9914	1.59
4	4	12	36	6	0.9872	1.67
6	2	18	54	12	0.9928	4.75
6	3	18	54	12	0.9891	5.12
6	4	18	54	12	0.9843	5.84
8	2	24	72	24	0.9896	9.85
8	3	24	72	24	0.9852	10.02
8	4	24	72	24	0.9815	10.15

TABLE VII
CROSS-DATASET PERFORMANCE ANALYSIS

Training Dataset	Testing Dataset	Logistic Regression	Random Forest	Proposed
UNSW-NB15	CIC-DDoS2019	0.88–0.91	0.92–0.95	**0.98–0.99**
UNSW-NB15	AppLayer DDoS	0.86–0.90	0.91–0.94	**0.97–0.99**
CIC-DDoS2019	UNSW-NB15	0.87–0.91	0.93–0.96	**0.98–0.99**
CIC-DDoS2019	AppLayer DDoS	0.89–0.92	0.94–0.96	**0.99**
AppLayer DDoS	UNSW-NB15	0.85–0.89	0.90–0.93	**0.97–0.98**
AppLayer DDoS	CIC-DDoS2019	0.88–0.91	0.92–0.95	**0.98–0.99**

where ρ_i denotes the utilization ratio of resource i, π_i represents the nominal power draw of the resource under full load, and τ_{inf} is the inference latency. The utilization ratios ρ_{CPU} and ρ_{GPU} are obtained from Table V, while $\tau_{\text{inf}} \in [1.0, 1.6]$ ms as reported in the experimental evaluation. For IoT gateway or edge-node deployment without GPU acceleration, the energy model reduces to $\mathcal{E}_{\text{edge}} = \rho_{\text{CPU}} \cdot \pi_{\text{CPU}} \cdot \tau_{\text{inf}}$, indicating further reduction in operational energy cost. The reported values show that $\mathcal{E}_{\text{inf}} < 1.5$ mJ across all datasets, confirming the suitability of the proposed framework for energy-constrained IoT environments shows in Table VI.

D. Cross-Dataset Validation and Attack-Type Analysis

A shared quantum latent space for cross-dataset evaluation validation is employed for hyperparameter improvement and incentive evaluation. Test results are provided in Table VII. Due to different feature spaces, UNSW-NB15 and CIC-DDoS2019 classifications cannot be used. SHO feature selection and 8-qubit quantum angle encoding transformed characteristics from each dataset into a shared low-dimensional latent space. XGBoost classifier is trained on UNSW-NB15 data and tested on Cic-DDoS2019 data without retraining using fixed-dimension quantum embedding. Define classes $C = \{DoS, Recon, Exploit, \ldots\} for attack - type accuracy$. Computing for each class $c \in C$:

$$\text{Accuracy}_c = \frac{\#\text{ correctly classified samples of class } c}{\#\text{ samples of class } c}. \tag{37}$$

For cross-dataset validation, denote training dataset $\mathcal{D}_{\text{train}}$ and testing dataset $\mathcal{D}_{\text{test}}$. Report $\text{F1}_{\text{test}}(\mathcal{D}_{\text{train}} \to \mathcal{D}_{\text{test}})$, and quantify domain shift as:

$$\Delta_{\text{F1}} = \text{F1}_{\text{val}}(\mathcal{D}_{\text{train}}) - \text{F1}_{\text{test}}(\mathcal{D}_{\text{test}}). \tag{38}$$

Cross-Dataset Evaluation is to verify generalizability, the model is trained on one dataset and tested on another. CIC-DDoS2019 tested UNSW-NB15 model, and vice versa. To ensure methodological rigor, a stratified k-fold cross-validation procedure ($k = 5$) is used, with each fold using non-overlapping training (70%), validation (15%), and test (15%) splits to prevent data leakage. Every iteration, one fold is tested, while the other four are divided into training (70%) and validation (15%) subsets. UnSW-NB15 and CIC-DDoS2019 have differing numbers of initial features and feature semantics, however direct transfer is not done in the raw feature space. Instead, using the same quantum encoding architecture, each dataset is preprocessed, normalized, and turned into a fixed-dimensional quantum latent representation.A shared latent feature space is used to evaluate cross-dataset transferability based on attack-discriminative statistical patterns rather than feature-name correspondence. The methodology may evaluate generalization across diverse intrusion detection datasets while minimizing feature-space incompatibilities. Performance is tested across folds, and robustness is evaluated using bootstrap resampling to generate 95% confidence intervals. The accuracy above 96.8% across diverse traffic distributions and attack characteristics, with precision and recall surpassing 95% and tight confidence intervals ($\pm$0.7–1.1%). The results show consistent generalization under domain change and reduce overfitting to a single dataset. DDoS, exploit, SYN, UDP, HTTP flood, and Slowloris hacks are obtained above 99% accurate with the QRL-SHO-XGBoost framework. In varied intrusion types and benchmark datasets, detection performance is resilient and balanced.

E. Ablation Study

Ablation study quantified the contributions of proposed learning framework components. Each variant eliminates or modifies one module while maintaining the whole model's training settings, hyperparameters, and assessment processes. Enhanced state representation, composite reward formulation, adaptive policy learning, and deep model architecture comprise the proposed framework. The ablation trials isolate each component's effect. To evaluate state informativeness, contextual and auxiliary elements are deleted. This adjustment reduces the agent's environmental dynamics observation. Performance drop is evaluated:

$$\Delta J_{\text{state}} = J_{\text{Full}} - J_{\text{V1}}, \tag{39}$$

where J denotes the average episodic return. The full reward function is defined as:

$$r_t = w_1 r_{\text{task}} + w_2 r_{\text{efficiency}} + w_3 r_{\text{penalty}}. \tag{40}$$

Another ablation variant replaces the quantum feature encoding stage with its conventional equivalent while obtaining the same SHO optimization, reinforcement learning

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

TABLE VIII
COMPONENT-WISE ABLATION ANALYSIS

Variant	Acc	F1-score	Final Return	Std. Dev.	Convergence
Proposed Model	99.21	99.19	245.8	6.3	1.2×10^5
No Quantum Encoding	98.62	98.55	238.4	7.1	1.3×10^5
Reduced State	98.11	98.03	221.4	9.7	1.8×10^5
No Reward Shaping	97.54	97.42	193.6	14.2	2.4×10^5
No Adaptation	95.86	95.71	118.9	5.1	–
Shallow Network	97.89	97.81	210.7	8.4	1.5×10^5
No Exploration	97.65	97.58	202.3	12.8	1.1×10^5

configuration, and XGBoost classifier to evaluate the quantum module. This comparison isolates quantum-enhanced feature representation. Quantum feature representation enhances policy learning and optimization efficiency shows in Table VIII, reducing the final return from 245.8 to 238.4. Variants V2–V6 examine reward shaping, adaptive learning, model capacity, and exploration. The static baseline proves adaptive learning is necessary, reward shaping removal degrades the most, and lower state representation slows convergence.Exploration removal speeds convergence but lowers return, while model capacity determines return. Quantum-enhanced feature encoding, state modeling, reward design, adaptive learning, and appropriate representational capacity appear to boost performance.

F. Security Analysis

The QRL-SHO-XGBoost intrusion detection system detects known and zero-day attacks across network and application layers, providing strong defense for IoT and cyber-physical environments. Reduced redundancy and noise on the feature vector and feature space increase attack surface and robustness against feature poisoning and evasion when selecting features using Sea-Horse Optimization (SHO). Quantum Reinforcement Learning (QRL) enables adaptive policy and hyperparameter optimization to respond to new attack patterns and idea drift. Quantum feature encoding captures robust, complex, non-linear cross-correlations from stealthy (low rate) attacks for more accurate classification. XGBoost's excellent classification performance and millisecond inference latency made it suitable for real-time use in resource-constrained IoT environments. LIME-based explainability boosts operational openness, forensic analysis, and confidence. The methodology has numerous benefits, however, it relies on simulated quantum devices and assumes training data is trusted, which could lead to training-time data poisoning with current NISQ hardware. However, the system is durable, adaptable, explainable, and generalizable across datasets, with apparent areas for boosting adversarial robustness and quantum scalability.

G. Computational Complexity (CC)

Let N be the population size, T the number of SHO iterations, d original features, and τ_{fit} the time to train/evaluate the classifier on one feature subset. The wrapper SHO complexity is approximately:

$$\mathcal{O}\big(N\,T\,\tau_{\text{fit}}(k)\big), \tag{41}$$

where $k = \mathbb{E}[|\mathcal{S}|]$ is the average selected feature count; $\tau_{\text{fit}}(k)$ typically grows superlinearly with k for tree ensembles. The quantum subroutines add overhead that depends on simulator vs real device; for simulators, worst-case classical simulation is $\mathcal{O}(2^{\zeta})$ for $\zeta = 8$ qubits, while real-device wall-time scaling depends on shot count n_{shots} and circuit runtime: $\mathcal{O}(n_{\text{shots}} \cdot \text{depth})$. The total computational complexity of the proposed QRL–SHO–XGBoost framework can be expressed as:

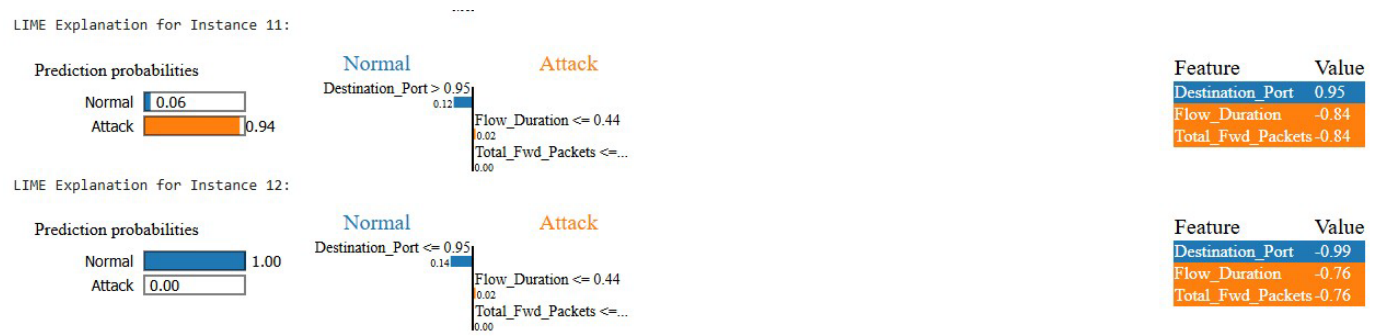
$$\begin{aligned}\mathcal{C}_{\text{Proposed}} \approx\ & \mathcal{O}(N \cdot T \cdot \tau_{\text{fit}}(k)) + \mathcal{O}(T_b \cdot n \cdot k \cdot \log n) \\ & + \mathcal{O}(E \cdot S \cdot A \cdot n_{\text{shots}} \cdot d_q \cdot \zeta),\end{aligned} \tag{42}$$

where N and T denote the SHO population size and iterations, $\tau_{\text{fit}}(k)$ the classifier fitting time on k selected features, E, S and A are the number of episodes, states, and actions in QRL, n_{shots} the number of circuit shots, d_q the circuit depth, ζ the number of qubits, T_b the number of boosting rounds, and n the number of training samples. For practical qubit sizes ($\zeta \leq 8$), the XGBoost component dominates; however, as qubits scale beyond 20, the quantum circuit evaluation overhead becomes significant due to exponential simulation costs.

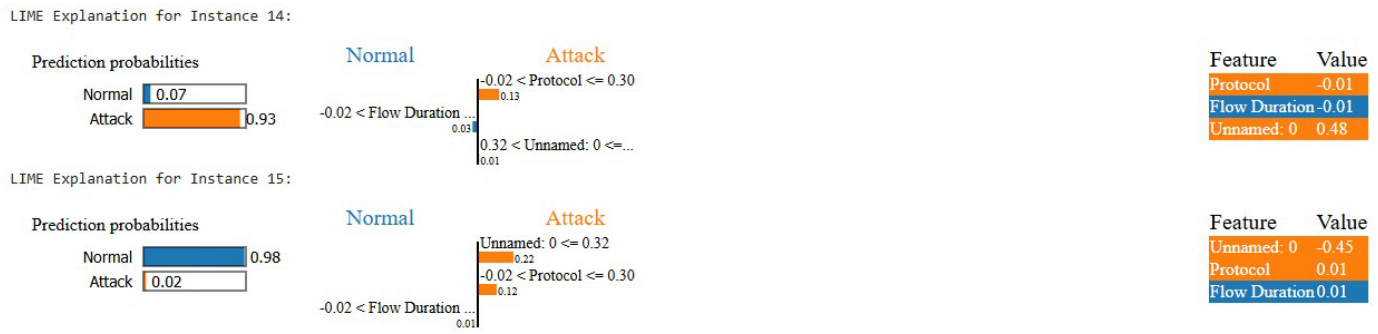
H. Explainability Using LIME

Explainability is vital for IDS to ensure transparency and trust in automated decisions. In our proposed SHO-QRL-XGBoost framework, we have employed LIME to provide input-state explanations for the predictions generated by the model. The identified features represent the highly influential features responsible for the final predictions in original feature space. LIME is responsible for providing approximate feature-level insights into the overall prediction process and helps in analyzing the attack detection behavior. The results are not interpreted as a direct mapping between network features and quantum measurement outcomes. For the interpretation of the prediction results of the proposed SHO-QRL-XGBoost framework, LIME-based local explanations were generated across 3 benchmark datasets with different set of features (necessary input features for LIME). For the Application DDoS dataset, Destination_Port, Flow_Duration, and Total_Fwd_Packets were identified as primary input features for contributing classifications as Normal or Attack, the primary classes to classify the predictions. In Fig. 5 (a), for a randomly chosen instance 11, LIME interpreted the prediction with 94% confidence as Attack (left side). In the middle, the LIME tabular provides the feature contributions with black bars and red bars, each pushing the LIME decision toward Normal and Attack respectively. The length of these bars represent the strength of the feature's contribution. Destination_Port is a highly influential feature and contributes to Normal class, whereas, Flow_Duration, and Total_Fwd_Packets contribute to Attack class. On the right side of the LIME interpretation represents the feature values of each feature, that are the normalized values of the highly influential feature for the particular instance. The 3 features Destination_Port, Flow_Duration, and Total_Fwd_Packets have feature values 0.95, −0.84, and −0.84 respectively. Similarly, considering instance 12, LIME predicted Normal with 100% confidence, with Destination_Port contributing towards Normal class and Flow_Duration, and Total_Fwd_Packets contributing towards Attack class. In

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.



(a) Application DDoS dataset



(b) CIC-DDoS Dataset



(c) UNSW Dataset

Fig. 5. LIME results on various datesets.

a similar way, for CIC-DDoS2019 dataset (Fig. 5 (b)), the contributing features provided as input features to LIME were Protocol, FlowDuration, and Unnamed (representing connection-related attributes). For UNSW-NB15 dataset (Fig. 5 (c)), the input features to LIME were IPV4_SRC_ADDR (source address), IPV4_DST_ADDR (destination address), and L4_SRC_PORT (Layer 4 source port). The explanations provided by LIME suggest local input-space interpretations of model's final predictions.

V. Conclusion and Future Work

The rapid growth of IoT networks has increased security vulnerabilities, challenging traditional IDS approaches that struggle with high-dimensional data, evolving attacks, and computational inefficiencies. We propose an advanced IDS framework integrating SHO, QRL, and XGBoost, evaluated on UNSW-NB15, CIC-DDoS2019, and AppLayer DDoS datasets. Experimental evaluation on the UNSW-NB15, CIC-DDoS2019, and AppLayer DDoS datasets demonstrated strong detection performance, achieving up to 99.21% accuracy and 99.19% F1-score while maintaining an average inference latency of approximately 1.2 ms per sample. Ablation studies further showed that the complete SHO-QRL-XGBoost framework outperformed XGBoost, SHO-XGBoost, and RL-XGBoost baselines, confirming the complementary contributions of feature optimization, adaptive learning, and quantum feature representation. In addition, cross-dataset experiments achieved accuracy exceeding 96% with narrow confidence intervals, indicating stable performance across heterogeneous intrusion detection datasets. LIME-based analysis highlighted the influence of key network traffic features on classification decisions, improving the interpretability of the proposed framework. Future work will focus on evaluating the framework on real quantum hardware, investigating larger quantum feature representations, improving cross-dataset generalization through stricter temporal and session-aware evaluation protocols, and extending the explainability analysis using complementary techniques such as SHAP and permutation-based feature importance. Additional

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.

research will also explore adversarial robustness and large-scale deployment in real-world IoT and cyber-physical system environments.

REFERENCES

[1] S. S. Gill and R. Buyya, "Transforming research with quantum computing," *J. Economy Technol.*, vol. 4, pp. 1–8, Feb. 2026.

[2] A. A. Khan, M. Driss, W. Boulila, G. A. Sampedro, S. Abbas, and C. Wechtaisong, "Privacy preserved and decentralized smartphone recommendation system," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 4617–4624, Feb. 2024.

[3] T. Zoppi, M. Gharib, M. Atif, and A. Bondavalli, "Meta-learning to improve unsupervised intrusion detection in cyber-physical systems," *ACM Trans. Cyber-Phys. Syst.*, vol. 5, no. 4, pp. 1–27, Oct. 2021.

[4] R. W. Anwer, M. Abrar, M. Ullah, A. Salam, and F. Ullah, "Advanced intrusion detection in the industrial Internet of Things using federated learning and LSTM models," *Ad Hoc Netw.*, vol. 178, Nov. 2025, Art. no. 103991.

[5] R. W. Anwer, M. Abrar, A. Salam, and F. Ullah, "TEAD: Trust-enhanced anomaly detection framework for intrusion detection in IoT-enabled wireless sensor networks (WSNs)," *Wireless Netw.*, vol. 31, no. 6, pp. 4179–4197, Aug. 2025.

[6] J. Li, X.-M. Li, G. Chen, X.-J. Peng, and H. Li, "Optimal tracking control for cyber-physical systems under mixed attacks via game-theoretical Q-learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 11944–11954, 2025.

[7] F. Ullah et al., "Lightweight multimedia anomaly and integrity detection for consumer IoT using knowledge distillation," *IEEE Trans. Consum. Electron.*, vol. 72, no. 1, pp. 2465–2475, Feb. 2026.

[8] I. A. Soomro, H. U. R. Khan, S. J. Hussain, Z. Ashraf, M. M. Alnfiai, and N. N. Alotaibi, "Lightweight privacy-preserving federated deep intrusion detection for industrial cyber-physical system," *J. Commun. Netw.*, vol. 26, no. 6, pp. 632–649, Dec. 2024.

[9] F. Ullah et al., "Deep trust: A novel framework for dynamic trust and reputation management in the Internet of Things (IoT)-based networks," *IEEE Access*, vol. 12, pp. 87407–87419, 2024.

[10] R. W. Anwar, M. Abrar, A. Salam, and F. Ullah, "Federated learning with LSTM for intrusion detection in IoT-based wireless sensor networks: A multi-dataset analysis," *PeerJ Comput. Sci.*, vol. 11, p. 2751, Feb. 2025.

[11] H. Asmat, F. Ullah, A. A. Khan, F. Ali, and M. I. Mohmand, "A novel caching framework for information-centric IoT using deep reinforcement proximal policy optimization," *Comput. Commun.*, vol. 241, Feb. 2025, Art. no. 108261.

[12] I. Singh, S. S. Gill, J. Sun, and J. Mol, "QAISim: A toolkit for modeling and simulation of AI in quantum cloud computing environments," *Cluster Comput.*, vol. 29, no. 2, p. 99, Apr. 2026.

[13] F. V. Massoli, L. Vadicamo, G. Amato, and F. Falchi, "A leap among quantum computing and quantum neural networks: A survey," *ACM Comput. Surv.*, vol. 55, no. 5, pp. 1–37, 2022.

[14] V. Rishiwal, U. Agarwal, M. Yadav, S. Tanwar, D. Garg, and M. Guizani, "A new alliance of machine learning and quantum computing: Concepts, attacks, and challenges in IoT networks," *IEEE Internet Things J.*, vol. 12, no. 12, pp. 18865–18886, Jun. 2025.

[15] M. Hdaib, S. Rajasegarar, and L. Pan, "Quantum deep learning-based anomaly detection for enhanced network security," *Quantum Mach. Intell.*, vol. 6, no. 1, p. 26, Jun. 2024.

[16] D. L. Marino, C. S. Wickramasinghe, C. Rieger, and M. Manic, "Self-supervised and interpretable anomaly detection using network transformers," *IEEE Trans. Ind. Informat.*, vol. 21, no. 5, pp. 4252–4261, May 2025.

[17] A. Kottahachchi Kankanamge Don and I. Khalil, "QRLaXAI: Quantum representation learning and explainable AI," *Quantum Mach. Intell.*, vol. 7, no. 1, pp. 1–22, Jun. 2025.

[18] M. Keshk, N. Koroniotis, N. Pham, N. Moustafa, B. Turnbull, and A. Y. Zomaya, "An explainable deep learning-enabled intrusion detection framework in IoT networks," *Inf. Sci.*, vol. 639, Aug. 2023, Art. no. 119000.

[19] U. Ahmed et al., "Explainable AI-based innovative hybrid ensemble model for intrusion detection," *J. Cloud Comput.*, vol. 13, no. 1, p. 150, Oct. 2024.

[20] F. Ullah et al., "Privacy-aware secure data auditing for cloud-based intelligence of things environment," *IEEE Internet Things J.*, vol. 12, no. 11, pp. 15288–15303, Jun. 2025.

[21] S. Zhao, T. Zhang, S. Ma, and M. Wang, "Sea-horse optimizer: A novel nature-inspired meta-heuristic for global optimization problems," *Appl. Intell.*, vol. 53, no. 10, pp. 11833–11860, May 2023.

[22] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 849–872, May 2019.

[23] A. Kadi, A. Selamnia, Z. A. E. Houda, H. Moudoud, B. Brik, and L. Khoukhi, "An in-depth comparative study of quantum-classical encoding methods for network intrusion detection," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 1129–1148, 2025.

[24] Y. Zhang, P. Tino, A. Leonardis, and K. Tang, "A survey on neural network interpretability," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 5, no. 5, pp. 726–742, Oct. 2021.

[25] C. Harvey, R. Yeung, and K. Meichanetzidis, "Sequence processing with quantum-inspired tensor networks," *Sci. Rep.*, vol. 15, no. 1, p. 7155, Feb. 2025.

[26] A. Khatun and M. Usman, "Quantum transfer learning with adversarial robustness for classification of high-resolution image datasets," *Adv. Quantum Technol.*, vol. 8, no. 1, Jan. 2025, Art. no. 2400268.

[27] W. Lu, Y. Lu, J. Li, A. Sigov, L. Ratkin, and L. A. Ivanov, "Quantum machine learning: Classifications, challenges, and solutions," *J. Ind. Inf. Integr.*, vol. 42, Nov. 2024, Art. no. 100736.

[28] M. Kalinin and V. Krundyshev, "Security intrusion detection using quantum machine learning techniques," *J. Comput. Virol. Hacking Techn.*, vol. 19, no. 1, pp. 125–136, Jun. 2022.

[29] A. Alomari and S. A. Kumar, "DEQSVC: Dimensionality reduction and encoding technique for quantum support vector classifier approach to detect DDoS attacks," *IEEE Access*, vol. 11, pp. 110570–110581, 2023.

[30] M. Y. Küçükkara, F. Atban, and C. Bayılmış, "Quantum-neural network model for platform independent DDoS attack classification in cyber security," *Adv. Quantum Technol.*, vol. 7, no. 10, Oct. 2024, Art. no. 2400084.

[31] M. Moll and L. Kunczik, "Comparing quantum hybrid reinforcement learning to classical methods," *Hum.-Intell. Syst. Integr.*, vol. 3, no. 1, pp. 15–23, Mar. 2021.

[32] D. S. Rajput and A. K. Upadhyay, "An innovative ensemble approach to detecting distributed denial of service attacks using machine learning," in *Proc. Int. Conf. Artif. Intell. Quantum Comput.-Based Sensor Appl. (ICAIQSA)*, Dec. 2024, pp. 1–6.

[33] T. H. Kim and S. Madhavi, "Quantum intrusion detection system using outlier analysis," *Sci. Rep.*, vol. 14, no. 1, p. 27114, Nov. 2024.

[34] M. Asif et al., "Intelligent two-phase dual authentication framework for Internet of Medical Things," *Sci. Rep.*, vol. 15, no. 1, p. 1760, Jan. 2025.

[35] C. Gong, W. Guan, H. Zhu, A. Gani, and H. Qi, "Network intrusion detection based on variational quantum convolution neural network," *J. Supercomput.*, vol. 80, no. 9, pp. 12743–12770, Jun. 2024.

[36] P. J. G. Lisboa, S. Saralajew, A. Vellido, R. Fernández-Domenech, and T. Villmann, "The coming of age of interpretable and explainable machine learning models," *Neurocomputing*, vol. 535, pp. 25–39, May 2023.

[37] A. Thuy and D. F. Benoit, "Explainability through uncertainty: Trustworthy decision-making with neural networks," *Eur. J. Oper. Res.*, vol. 317, no. 2, pp. 330–340, Sep. 2024.

[38] H. Wang, Q. Liang, J. T. Hancock, and T. M. Khoshgoftaar, "Feature selection strategies: A comparative analysis of SHAP-value and importance-based methods," *J. Big Data*, vol. 11, no. 1, p. 44, Mar. 2024.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on November 20,2026 at 03:24:49 UTC from IEEE Xplore. Restrictions apply.