EQAI: Explainable Quantum-Empowered Antispoofing Intelligence for Trustworthy Connected Autonomous Vehicles Communication

Simeon Okechukwu Ajakwe, *Senior Member, IEEE*, and Dong-Seong Kim, *Senior Member, IEEE*

***Abstract*—The increasing complexity of connected autonomous vehicles (CAVs) introduces new security challenges in the Internet of Vehicles (IoV), where traditional detection models struggle with real-time interpretability, scalability, and robustness against spoofing attacks. This article proposes an explainable quantum-empowered antispoofing intelligence (EQAI) framework that fuses quantum-enhanced learning with interpretable trust inference for securing vehicular communications. The EQAI model employs an eight-qubit variational quantum circuit (VQC) integrated with lightweight classical layers and explainability modules based on Shapley additive explanation (SHAP) and local interpretable model-agnostic explanation (LIME). Using Canadian Institute for Cybersecurity Internet of Vehicles 2024 Dataset (CICIoV2024), the framework achieves a detection accuracy of 92.85%, a Class 3 $F1$-score of 82.1%, and a low false alarm rate (FAR $\leq$ 0.026), outperforming existing machine learning (ML), blockchain (BC)-based, and federated learning (FL) approaches. Its compact design—with only 4900 trainable parameters and ~19.5 k floating point operations per second (FLOPs)—demonstrates real-time deployability and energy efficiency at the network edge. Moreover, LIME and SHAP analyses reveal transparent feature-level reasoning, enhancing operator trust and system explainability. The proposed EQAI architecture thus establishes a scalable, interpretable, and quantum-empowered foundation for secure, trustworthy, and intelligent vehicular networks, advancing toward resilient IoV and next-generation cybercognitive mobility ecosystems.**

***Index Terms*—Connected autonomous vehicles (CAVs), explainable AI (XAI), quantum information theory, spoofing attacks, trustworthy communication, vehicle-to-everything (V2X).**

NOMENCLATURE

AI	Artificial intelligence.
AUC	Area under the receiver operating characteristic curve.
BC	Blockchain.
CAV	Connected autonomous vehicles.
CICIoV2024	Canadian Institute for Cybersecurity Internet of Vehicles 2024 Dataset.
DL	Deep learning.
DT	Digital twin.
EQAI	Explainable quantum-empowered antispoofing intelligence.
FAR	False alarm rate.
FLOPs	Floating point operations per second.
FL	Federated learning.
IoV	Internet of Vehicles.
IDS	Intrusion detection system.
LIME	Local interpretable model-agnostic explanations (local explainability).
ML	Machine learning.
MLP	Multilayer perceptron.
MD	Missed detection rate.
PQC	Postquantum cryptography.
QAI	Quantum AI.
QML	Quantum ML.
QNN	Quantum neural network.
QKD	Quantum key distribution.
SHAP	Shapley additive explanations (global explainability).
TAD	Trust assessment and decision-making.
V2X	Vehicle-to-everything.
VQC	Variational quantum circuit.
XAI	Explainable AI.
q	Number of qubits used in the quantum circuit.
L	Number of variational quantum layers (repetitions).
Θ_Q	Trainable quantum parameters in the EQAI model.
Θ_C	Trainable classical parameters in the EQAI model.
Θ	Total trainable parameters ($\Theta = \Theta_Q + \Theta_C$).
α	Learning rate used during model optimization.
T	Trust score derived from explainable inference.

Received 21 October 2025; revised 24 November 2025; accepted 25 November 2025. Date of publication 28 November 2025; date of current version 9 March 2026. This work was supported in part by the Innovative Human Resource Development for Local Intellectualization program through the Institute of IITP grant funded by the Korean government [Ministry of Science and ICT (MSIT)] under Grant IITP-2024-RS-2020-II201612, 33%, in part by the Priority Research Centers Program through the NRF funded by the Ministry of Education, Science and Technology (MEST) under Grant 2018R1A6A1A03024003, 33%, and in part by the IITP (Institute of Information and Communications Technology Planning and Evaluation)-ICAN [ICT Challenge and Advanced Network of Human Resources Development (HRD Korea)] grant funded by the Korea government [Ministry of Science and Information and communications technology (ICT)] under Grant IITP-2024-2022-00156394, 34%. *(Corresponding author: Dong-Seong Kim.)*

Simeon Okechukwu Ajakwe was with the ICT Convergence Research Center, Kumoh National Institute of Technology, Gumi 39177, South Korea. He is now with the Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, South Korea (e-mail: simeonajlove@gmail.com).

Dong-Seong Kim is with the Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, South Korea (e-mail: dskim@kumoh.ac.kr).

Digital Object Identifier 10.1109/JIOT.2025.3638378

F1	Weighted $F1$-score for spoofing classification performance.
Acc	Detection accuracy (%).
t_{lat}	Average model inference latency (ms).
E_q	Quantum embedding energy (used in variational encoding).
P_{legit}	Probability of legitimate message classification.
P_{spoof}	Probability of spoofed message classification.

I. Introduction

CONNECTED autonomous vehicles (CAVs) are transforming modern transportation by enabling real-time navigation, reduced human error, and improved traffic efficiency through vehicle-to-everything (V2X) communication [1]. However, this connectivity introduces substantial cybersecurity risks, particularly spoofing attacks, where adversaries impersonate legitimate entities to inject falsified messages into the network. Such attacks can lead to incorrect routing decisions, vehicular collisions, or even fatal outcomes [2]. Ensuring the integrity, authenticity, and trustworthiness of CAV communications is, therefore, critical to the safe deployment of intelligent transportation systems.

Traditional security solutions, such as cryptographic authentication and data encryption, offer basic safeguards but struggle to detect subtle or adaptive spoofing techniques [3]. Moreover, classical ML approaches, though widely used, often lack robustness to adversarial attacks and fail to provide transparency in decision-making [4]. For example, random forest (RF) models and FL architectures have been proposed for intrusion and spoofing detection in vehicular networks [5]. However, these methods are often limited by low adaptability, lack of interpretability, and computational inefficiencies in handling high-dimensional, real-time CAV data as summarized in Table I.

To address these challenges, we propose a novel framework, explainable quantum-empowered antispoofing intelligence (EQAI), that integrates quantum information theory with XAI. QAI leverages principles of quantum mechanics, such as superposition and entanglement, to enable high-capacity parallel processing and enhanced pattern recognition [14]. On the other hand, XAI focuses on making AI decision processes transparent and interpretable, using tools such as SHAP and LIME [6].

A. Motivation for QAI Models in Spoofing Detection

Spoofing attacks in CAVs often exhibit subtle variations in signal parameters such as phase, amplitude, delay, and frequency, making them difficult to detect using traditional ML models. Classical approaches typically operate in a linear or shallow feature space, which limits their ability to distinguish between legitimate and adversarial signals, especially when noise or adversarial perturbations are present. For instance, in aerial-based logistics using drones, curtailing such intrusion via classical AI-driven approaches has remained a daunting issue [15]. QML models, in contrast, harness the principles of quantum mechanics—superposition, entanglement, and interference, to encode and process data in exponentially larger Hilbert spaces [16]. This allows QML to explore complex, high-dimensional data distributions in parallel, making it highly effective at capturing the nuanced patterns associated with spoofing attacks. In addition, quantum circuits naturally model uncertainty and probabilistic behavior [17], which aligns well with the stochastic nature of over-the-air vehicular communication. By exploiting this quantum parallelism, the EQAI framework enhances detection sensitivity and reduces the likelihood of false negatives, especially in zero-day spoofing scenarios.

The proposed EQAI framework comprises three core modules: 1) the quantum-enhanced spoofing threat (QUEST) detection module; 2) the explainable hybrid AI ($\hat{x}$AI) module; and 3) the trust assessment and decision-making (TAD) module. Together, these modules provide robust spoofing detection, enhance the interpretability of decisions, and assess message trustworthiness in real time. To the best of our knowledge, no work has combined QAI and XAI for spoofing detection in CAVs.

The key contributions of this article are as follows.

1) We propose EQAI, a hybrid quantum-classical framework for detecting and mitigating spoofing attacks in CAV communication.
2) We develop a quantum-enhanced learning model integrated with XAI tools (SHAP and LIME) to provide both accuracy and explainability in threat detection.
3) We evaluate EQAI using a publicly available GPS spoofing dataset and compare its performance against classical ML baselines in terms of accuracy, latency, communication overhead, and robustness.

The rest of the article is organized as follows. Section II explores previous approaches and provides insights about existing gaps addressed. Section III presents the system architecture and methodology. Section IV discusses experimental results and performance evaluation. Section V concludes the study with future directions.

II. Related Works

Spoofing detection and trust management in CAV networks have attracted significant research interest due to the increasing reliance on GPS and intervehicle communication for critical navigation decisions. Traditional trust-based approaches in vehicular networks, such as those surveyed in [18], compute node trust using historical behavior or context metrics. Although intuitive, they rarely incorporate real-time learning or model-driven explanations. XAI has been introduced in safety-critical AI systems, but has only recently been applied to spoofing defense. Most existing work either lacks real-time applicability or focuses solely on feature attribution without integrating it into a trust-based decision pipeline. Table I summarizes notable recent approaches addressing this problem using a variety of cryptographic, statistical, and learning-based methods.

Cryptographic schemes such as navigation message authentication (NMA) [19] enhance GNSS integrity using symmetric

TABLE I
COMPARATIVE ANALYSIS OF PROPOSED EQAI WITH RELATED WORKS

Reference	Target Area / Technique	Contributions	Limitations	Gaps Addressed by EQAI
[6] Classical XAI (RF)	Intrusion / spoofing detection using Random Forest with explainability	Introduced explainable AI for IoV security	Vulnerable to sophisticated spoofing; limited adaptability	EQAI integrates quantum-enhanced detection with SHAP/LIME explainability for robustness against advanced (Class 3) attacks.
[5] Classical Federated Learning (KNN)	FL for GPS spoofing detection	Lightweight collaborative learning model	Detects only simple spoofing; vulnerable to poisoning	EQAI provides stronger resilience to adversarial spoofing and dynamic trust scoring beyond FL aggregation.
[3] Homomorphic Encryption + FL (DL)	Privacy-preserving spoofing detection	Combines FL with HE for secure training	High computational cost; unsuitable for real-time vehicular context	EQAI achieves moderate overhead with hybrid quantum-classical circuits, ensuring onboard feasibility.
[7] Blockchain + QAI	Blockchain-secured spoofing detection	Blockchain-based QAI integration	High communication/-computation overhead	EQAI uses lightweight trust-aware decision (TAD) with reduced communication overhead (0.003 MB).
[8] PQC + Blockchain	Quantum-resilient blockchain framework for CAV communication	Introduces PQC for quantum-safe V2X trust	Blockchain latency and heavy consensus delay	EQAI complements PQC-secured frameworks by adding explainable trust scoring and low-latency inference.
[9] PQC for In-vehicle Security	Post-quantum cryptography (Kyber) for CAN bus protection	Strengthens intra-vehicle communication security	Focuses on intra-vehicle links only; lacks trust scoring and explainability	EQAI extends protection to inter-vehicle spoofing with interpretable trust metrics.
[10] QuantumShield-BC (Blockchain + PQC + QKD)	Quantum-secured blockchain with Byzantine consensus	Combines PQC + QKD + blockchain consensus for secure V2X	High complexity and deployment overhead in real-time systems	EQAI offers a lightweight hybrid model optimized for real-time V2X, balancing complexity and trust.
[11] AIoT + Digital Twin + DRL	AIoT-enabled digital twins with DRL-based trajectory planning	Supports adaptive trajectory optimization in CAVs	Focuses on trajectory; lacks spoofing resilience or message trust validation	EQAI complements trajectory planning by ensuring secure, explainable, and trustworthy message validation.
[12] Digital Twin AI-based Architecture	Data-driven DT for AV co-design	Supports simulation and system optimization	No focus on spoofing or security; limited real-time explainability	EQAI integrates DT compatibility with real-time quantum-enhanced spoofing detection.
[13] RL-based Motion Planning Survey	Survey of reinforcement learning for autonomous driving	Highlights DRL utility for trajectory planning	Does not address communication trust or adversarial spoofing threats	EQAI ensures secure V2X data as a foundation for safe DRL-driven trajectory planning.
Proposed EQAI (this work)	Hybrid Quantum–Classical + Explainable Trust Engine	Robust spoofing detection (Class 3), interpretable SHAP/LIME insights, and lightweight TAD-based trust scoring	Slightly lower average accuracy vs. MLP (trade-off for interpretability)	Addresses gaps in (i) explainability, (ii) robustness to sophisticated spoofing, (iii) real-time trust scoring, and (iv) lightweight deployment for CAVs.

time-based keying (e.g., TESLA) or dedicated L1-SAIF channels. While secure, these methods require GNSS infrastructure modifications and often introduce latency due to cryptographic verification. Moreover, they lack adaptability to novel spoofing vectors or context-aware trust evaluation. Also, threshold-based approaches like the one proposed by Leinmüller et al. [20] apply plausibility heuristics based on mobility grade, location continuity, or speed bounds. These solutions are lightweight and interpretable, but prone to false negatives when attackers craft spoofing patterns that mimic realistic trajectories.

Furthermore, learning-based strategies have shown improved adaptability. For instance, Xu et al. [21] proposed an FL IDS using LSTM models with differential privacy. Although effective against a broad class of anomalies, the method lacks dedicated GPS spoofing context and explainability mechanisms, making it difficult to diagnose or justify alarms. Similarly, hybrid frameworks proposed by Ayaz et al. [22] integrated FL with BC consensus to enhance trust propagation in message dissemination. However, their approach incurs high latency and does not explicitly support per-message spoofing defense or interpretability, as it does not provide actionable insights into why a message is flagged as malicious. Recent efforts in self-supervised learning, such as the work by Liu and Papadimitratos [23], propose anomaly-aware federated GNSS spoofing detection. This method improves resilience to evolving spoofing attacks through self-generated labels, but still lacks interpretable outputs and real-time trust decision frameworks.

Recent efforts have also explored quantum-resilient and BC-enabled secure communication frameworks for CAVs. For example, Aslam et al. [8] proposed a BC-assisted architecture that integrates PQC to safeguard vehicular message exchanges against quantum adversaries. Similarly, Castiglione and Elia [9] demonstrated the feasibility of securing in-vehicle controller area network (CAN) communications using CRYSTALS-Kyber, a NIST-standardized PQC primitive, thereby strengthening the resilience of vehicular systems against quantum attacks. Furthermore, Reddy et al. [10] introduced a quantum-secured BC framework (QuantumShield-BC) that combines PQC, QKD, and Byzantine consensus to enhance trust in vehicular communication networks. Complementary work by Nsour [24] proposed a lightweight onboard PQC model for autonomous vehicles, validating the practical deployment of quantum-resilient cryptography in CAV platforms. These contributions highlight that PQC-enabled BC systems are emerging as a critical enabler of next-generation secure V2X communications.

On the other hand, in the domain of AIoT-driven DTs and trajectory planning, Mohd Aslam et al. [11] developed an AIoT-enabled DT framework for CAVs, employing deep reinforcement learning (DRL) to optimize real-time trajectory planning under dynamic traffic conditions. Babaei et al. [12] further extended this vision by proposing a data-driven DT AI architecture for autonomous vehicles, supporting simulation-based co-design and decision optimization. In addition, Li et al. [13] surveyed reinforcement learning approaches to motion planning, providing insights into the integration of DRL for adaptive and safe vehicular navigation. Kabir

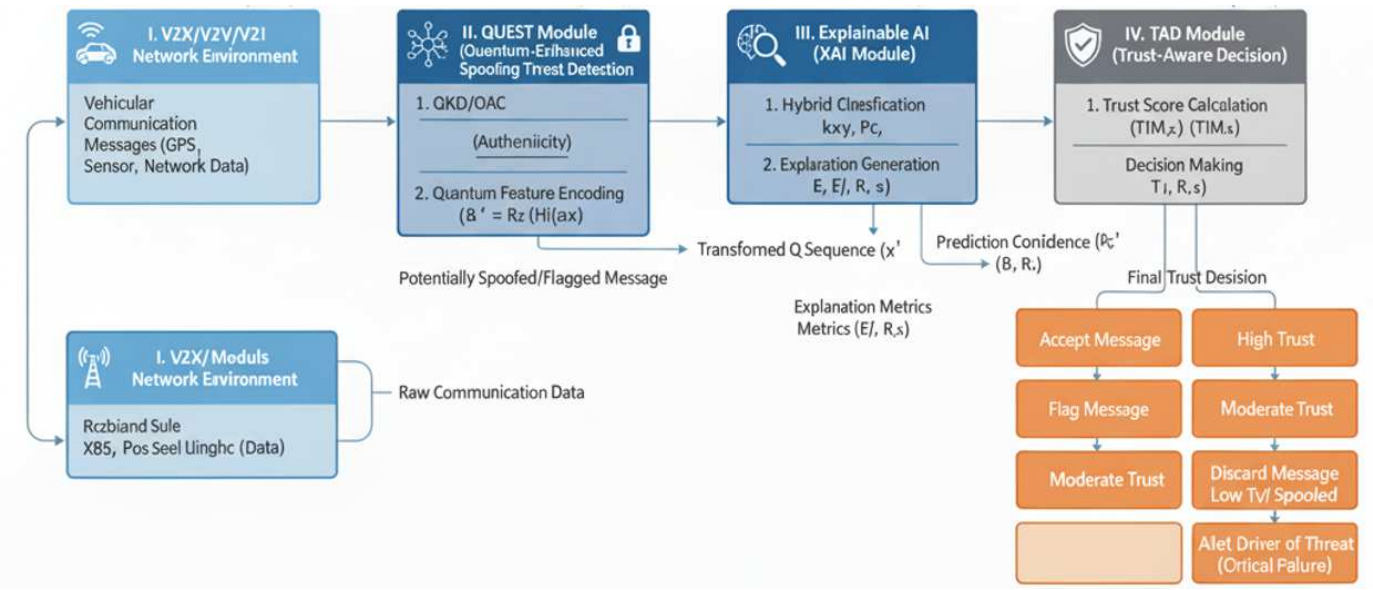

Fig. 1. EQAI system design/architecture for trustworthy V2X communication highlighting: 1) QUEST module, 2) XAI module, and 3) TAD module.

et al. [25] also presented a comprehensive survey on DT technologies for vehicular prototyping, underscoring the synergy between AIoT, DRL, and digital replicas for future CAV ecosystems. Together, these works demonstrate the shift toward intelligent, explainable, and simulation-augmented frameworks that complement our proposed EQAI-based trust model.

Despite the above advancements, current solutions often suffer from a lack of model explainability, making it difficult to interpret spoofing predictions or build user/operator trust. Also, there is an inability of these approaches to assess message-level trust dynamically in real time, which is critical for vehicular actionability. Moreover, they have limited robustness against sophisticated or adversarial spoofing strategies that mimic plausible mobility behaviors. Finally, previous approaches have heavy computation or communication overhead (e.g., BC-based FL), reducing real-time feasibility as captured in Table I. To address these limitations, the EQAI framework provides the following novelties.

1) *Quantum-Enhanced AI:* We incorporate quantum encoding to increase feature separation in Hilbert space, improving spoofing detection accuracy, especially for advanced attacks.
2) *Explainable Trust Assessment:* We fuse SHAP and LIME to achieve model-level explainability with classification decisions to provide interpretable feature attributions. This insight is further combined in a TAD module to derive real-time, actionable trust scores.
3) *Per-Class Performance Optimization:* Unlike closed box ML models, EQAI is optimized for robustness on complex spoofing classes (e.g., Class 3), achieving statistically significant gains with acceptable tradeoffs in other classes.
4) *Lightweight Execution:* The model is designed for real-time execution with moderate computational overhead and explainability costs, making it suitable for onboard vehicular deployment.

Overall, EQAI fills a critical gap by offering a trustworthy, interpretable, and quantum-assisted framework that supports secure and context-aware communication decisions in vehicular networks. By combining explainability, robustness, and real-time inference, EQAI represents a significant step toward trustworthy V2X communication in future CAV ecosystems.

III. EQAI System Design and Methodology for Trustworthy Vehicular Communication

The EQAI system is a layered, real-time trust pipeline meticulously designed to secure vehicle-to-vehicle (V2V), V2X, and V2I communications against sophisticated GPS spoofing attacks by embedding trustworthiness into the core decision-making process. The system's effectiveness stems from the seamless, low-latency interaction of its three principal modules. The proposed EQAI framework comprises three main components: 1) QUEST detection module; 2) explainable hybrid AI ($\widehat{x}$AI) module; and 3) trust assessment and decision-making (TAD) module, as seen in Fig. 1.

The process initiates with the QUEST module, which functions as the robust, quantum-secure front-end. This module first ensures message authenticity using QKD and quantum authentication codes (QACs). Crucially, it then employs a quantum feature encoding technique, transforming classical GPS and sensor data into a quantum-enhanced feature vector ($\mathbf{x}'$). This quantum representation utilizes the principles of Hilbert space for superior feature separation, making even subtle spoofing artifacts significantly more discernible than in classical feature spaces. The enhanced vector ($\mathbf{x}'$) is subsequently passed to the explainable AI ($\widehat{x}$AI) module. This module leverages a hybrid quantum-classical ML model to perform spoofing classification and generate a prediction confidence (P_c). Simultaneously, the XAI engine utilizes interpretation methodologies, such as SHAP and LIME, to produce

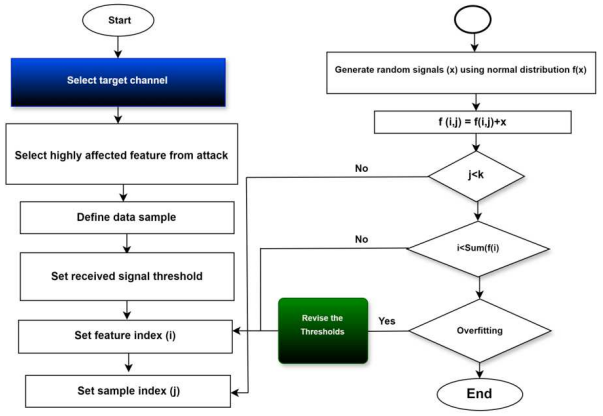

Fig. 2. Flowchart highlighting the steps to assess the status and category of a spoofed data as simplistic, intermediate, or sophisticated attack.

explanation metrics (E_f, R_s). These metrics detail why the model classified the message in a particular way, providing essential transparency that is vital for establishing trust in the system's output. Finally, the trust-aware decision (TAD) module integrates the classification confidence (P_c) and the explanation metrics (E_f, R_s). It calculates a dynamic and interpretable trust score $(T(M_t))$ for the vehicular message. By comparing this score against predefined trust thresholds, the TAD module makes the final, auditable trust decision—to `Accept`, `Flag`, or `Discard` the message—thereby ensuring that only verified, trustworthy information is acted upon by the CAV. This comprehensive, three-module interaction provides both unparalleled spoofing detection and the necessary explainability to guarantee security and trustworthiness in mission-critical vehicular communication.

A. QUEST Detection

Unlike classical information, the QUEST component takes advantage of the superposition and entanglement properties of quantum information to detect subtle deviations caused by spoofing attacks. QUEST involves (four) processes to intercept spoofing attacks: 1) QKD by establishing a shared secret key between communicating vehicles for authentication; 2) generating a QAC using the shared key and appending it to the message; 3) Transmitting the message with the QAC as a sequence of quantum states; and 4) verifying the QAC by measuring the quantum states using the shared secret key. If QAC is corrupted beyond a tolerable threshold, the message is flagged as potentially spoofed.

To assess the level of spoofed communication on CAV, we adopted a procedure as summarized in the flowchart in Fig. 2.

In this study, the spoofing attacks are categorized as simplistic, intermediate, and sophisticated based on the degree of the GPS and transmission power deviation over-the-air and the reception point.

B. Spoofing Signal Model

In a typical CAV communication scenario, the received GPS signal can be mathematically represented as

$$S(t) = A_s D(t-\tau_s)\, x(t-\tau_s) \cos\left(2\pi (f_l + f_d)\, t + \phi\right) \tag{1}$$

where

$S(t)$	received legitimate signal;
A_s	signal amplitude;
$D(t)$	digital data bit stream;
$x(t)$	spreading code;
τ_s	code phase delay;
f_l	carrier frequency;
f_d	Doppler frequency shift;
ϕ	carrier phase offset.

In the case of a spoofing attack, an adversary transmits N falsified signals that emulate legitimate transmissions, defined as

$$S_s(t) = \sum_{n=1}^{N} A'_s D'_s\left(t-\tau'_s\right) x\left(t-\tau'_s\right) \cos\left(2\pi\left(f_l + f'_d\right) t + \phi'\right) \tag{2}$$

where

$S_s(t)$	spoofed composite signal;
A'_s	amplitude of the spoofed signal;
$D'_s(t)$	spoofed data bit stream;
τ'_s	spoofed code phase delay;
f'_d	spoofed Doppler frequency shift;
ϕ'	spoofed carrier phase offset.

These spoofed signals aim to mimic the structure of legitimate signals but deviate in timing, frequency, and phase, which can be detected through the quantum-enhanced spoofing detection module. Thus, for a simplistic spoofing attack, $A_{s'} > A_s$ which increases the value of C/NO. For an intermediate spoofing attack, there is a distortion and an increase in delay and phase lock, which affects synchronization of code phase (τ') and Doppler (f'_d). Finally, for the sophisticated spoofing attack, the distortions and changes in correlation loops of peak-to-next-peak ratio in multiple parallel channels are considered using a quadrature accumulation shift correlator (PQP).

C. Explainable Hybrid AI Module ($\widehat{x}$ai)

The $\widehat{x}$AI module complements the quantum-powered QUEST engine by providing transparency into the spoofing detection decisions. The $\widehat{x}$AI module enhances the interpretability of the spoofing detection process in CAV environments. This section describes how classical vehicular communication data is encoded into quantum states, processed through a VQC, and integrated with classical neural networks to form a hybrid model.

D. Dataset Representation

The dataset is a collection of labeled feature vectors

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N} \tag{3}$$

where $x_i \in \mathbb{R}^n$ is a real-valued vector representing signal characteristics (e.g., CNO, PD, and CP), and $y_i \in \{0, 1, 2, 3\}$

represents message legitimacy (legitimate or spoofed message types).

E. Quantum Feature Encoding

To encode classical features into quantum states, we apply a parameterized quantum circuit $U_\phi(x')$ that transforms the input vector $x' \in \mathbb{R}^n$ into a quantum state over q qubits. For simplicity and compatibility with variational circuits, we choose $q = n$ and apply a single-qubit rotation encoding followed by a Hadamard gate. To convert classical input features into quantum states, we use

$$U_\phi\left(x'\right) = \bigotimes_{i=1}^{q} R_z\left(x_i'\right) H. \tag{4}$$

Here, $R_z(x_i')$ rotates each qubit about the Z-axis by the feature value x_i', injecting phase information. The Hadamard gate H places the qubit in superposition, enabling parallel processing in the Hilbert space. This design is ideal for encoding continuous-valued spoofing characteristics (e.g., PD, CP, and CNO) into quantum states. This combination enables the quantum circuit to capture both amplitude and phase-based characteristics of the input data, which is essential for identifying subtle anomalies in spoofed signals.

F. Variational Quantum Circuit

The encoded quantum state is passed through a trainable VQC

$$V\left(\Theta\right) = V_L\left(\theta^{(L)}\right) \cdots V_2\left(\theta^{(2)}\right) V_1\left(\theta^{(1)}\right). \tag{5}$$

Each V_l consists of entangling gates and parameterized single-qubit rotations with trainable parameters $\theta^{(l)}$ and L layers of entangling gates.

G. QNN Output

The output is obtained by measuring the expectation value of Pauli-Z observables on the final state

$$q_{\text{out}}\left(x';\Theta\right) = \left\langle 0 \middle| U_\phi^\dagger(x') V^\dagger\left(\Theta\right) O V\left(\Theta\right) U_\phi(x') \middle| 0 \right\rangle. \tag{6}$$

This measurement collapses the quantum state to a classical vector for further processing.

H. Classical Neural Network and Prediction

The output vector z from the quantum circuit is fed into a classical neural network $g(z; W)$ to generate final predictions

$$\hat{y} = g\left(\text{Process}\left(q_{\text{out}}\left(x';\Theta\right)\right); W\right). \tag{7}$$

Here, Process$(\cdot)$ is an optional transformation layer, and W is the weights of the classical network.

I. Loss Function and Optimization

Training is based on minimizing the categorical cross-entropy

$$\mathcal{L}\left(\hat{y}, y\right) = -\sum_{i=1}^{C} y_i \log\left(\hat{y}_i\right). \tag{8}$$

Parameters are updated using the Adam optimizer

$$\Theta \leftarrow \Theta - \alpha \cdot \nabla_\Theta \mathcal{L}, \quad W \leftarrow W - \alpha \cdot \nabla_W \mathcal{L}.$$

J. Hybrid Quantum-Classical Circuit Diagram

Fig. 3 is a visual representation of the hybrid EQAI model pipeline, highlighting the integration of quantum and classical layers for explainable and trustworthy spoofing attack detection. On the other hand, Fig. 4 captures the quantum circuit for four features (x_1–x_4) encoded using R_z and H, followed by a variational block $V(\Theta)$ and measurement to produce classical outputs z_i. A classical neural network then processes these outputs.

Table III captures the structural and training parameters of the proposed EQAI hybrid quantum model compared with its classical MLP counterpart. Bayesian optimization and the Adam optimizer with a learning rate ($\alpha = 0.01$) were adopted to achieve better model convergence with noisy gradients, with an appropriate batch size and epoch. While the classical MLP requires a larger parameter space to achieve comparable accuracy, the hybrid eight-qubit architecture attains similar or better spoofing detection performance with only 4900 trainable parameters, lower FLOPs (~19.5 k), and reduced latency and training overhead. The integration of quantum layers enhances robustness and nonlinear feature extraction by exploiting superposition and entanglement, enabling the model to capture subtle temporal–spatial correlations in vehicular signals that classical networks often overlook—thereby improving detection reliability against complex spoofing attacks in dynamic vehicular networks. This underscores the efficiency and robustness advantages of EQAI over the classical MLP in terms of model size, FLOPs, and enhanced nonlinear feature capture for reliable spoofing attack detection in vehicular networks.

K. TAD Module for CAV Communication

In CAVs, where split-second decisions govern safety, merely detecting spoofed messages is insufficient. Therefore, the EQAI framework incorporates a TAD module that performs two critical tasks: 1) evaluates the trustworthiness of each incoming message; and 2) triggers appropriate action (accept, flag, and reject) based on XAI insights. For the TAD trust computation pipeline, after a message passes through the QUEST and explainable hybrid AI ($\widehat{x}$AI) modules, the TAD module fuses the classification probability, feature attributions (SHAP or LIME), and historical sender reputation to compute a dynamic trust score $T \in [0, 1]$ as defined in the following equation:

$$T = \lambda_1 \cdot P_c + \lambda_2 \cdot E_f + \lambda_3 \cdot R_s \tag{9}$$

where P_c = classification confidence (softmax output for legitimate class); E_f = aggregated explanation fidelity (SHAP agreement with expected benign feature values); R_s = reputation of sender from historical logs; and λ_1, λ_2, and λ_3 = tunable weights based on context.

Regarding the trust score, $T \in [0, 1]$ as defined in (9) is required because T functions as a probability-like trust indicator used for threshold-based classification in the TAD module. Each component in

$$T = \lambda_1 P_c + \lambda_2 E_f + \lambda_3 R_s \tag{10}$$

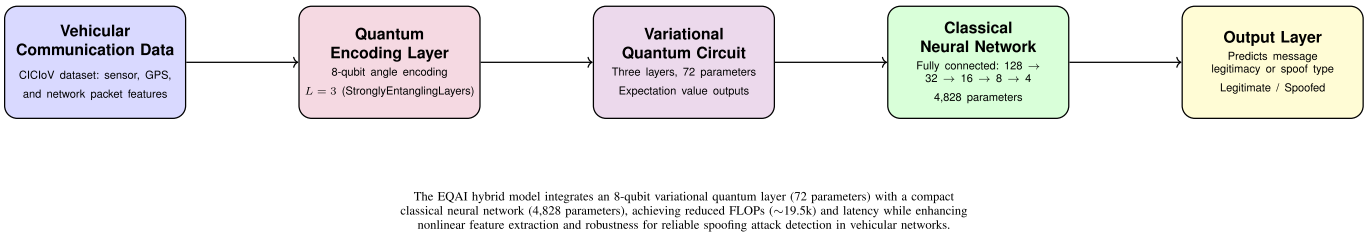

Fig. 3. EQAI hybrid quantum model pipeline illustrating the integration of quantum and classical layers for explainable spoofing attack detection in vehicular networks.

TABLE II
SUMMARY OF EQAI ARCHITECTURE PARAMETERS SHOWING ITS EFFICIENCY AND ROBUSTNESS ADVANTAGES OVER THE CLASSICAL MLP

Parameter	**Value & Description**
Number of qubits (q)	8. Physical qubits encoding classical features into quantum states.
Quantum layers (L)	3. Variational layers implemented using `StronglyEntanglingLayers`.
Quantum trainable parameters (Θ_Q)	72. Computed as $3 \times q \times 3 = 72$.
Classical layer configuration	Fully connected: $128 \to 32 \to 16 \to 8 \to 4$.
Classical trainable parameters (Θ_C)	Exact sum of layer parameters: (128×32+32) + (32×16+16) + (16×8+8) + (8×4+4) = 4,828.
Total trainable parameters (Θ)	$\Theta_Q + \Theta_C = 4,900$.
Optimizer	Adam optimizer (`lr=0.01`).
Learning rate (α)	0.01.
Batch size	32.
Epochs	50.
Frameworks	PennyLane (qml) for quantum circuit simulation and PyTorch for classical layers.
Explainability methods	SHAP and LIME for global and local interpretability.
Approx. FLOPs (forward pass)	∼19,500. Based on hybrid 8-qubit architecture estimation.

TABLE III
COMPARATIVE PERFORMANCE OF EQAI AGAINST CLASSICAL SOTA AI MODELS

Metrics/ Models	Acc (%)	Pr (%)	Rc (%)	F1 (%)	MSE (Loss)	False Alarm Rate (FAR) per Class				Missed Detections (MD) per Class				AUC-ROC per Class			
						0	1	2	3	0	1	2	3	0	1	2	3
RF	99.27	99.27	99.27	99.27	0.3150	0.0168	0.0021	0.0003	0.0014	0.0046	0.0269	0.0030	0.0244	99.90	99.81	98.59	98.32
LR	77.59	60.82	77.59	67.87	1.0000	0.9980	0.0000	0.0006	0.0014	0.0008	1.0000	0.9995	1.0000	63.67	62.67	62.15	61.50
XGBoost	92.61	92.55	92.61	92.56	0.2624	0.1556	0.015	0.0283	0.0037	0.0347	0.2183	0.2996	0.0815	99.13	98.15	98.10	98.05
SVM	75.00	75.00	75.00	75.00	0.2500	0.2500	0.2500	-	-	0.2500	0.2500	-	-	92.86	92.50	92.25	91.00
MLP	92.36	92.88	92.36	92.39	0.2795	0.1429	0.0064	0.0386	0.0039	0.0395	0.3424	0.1339	0.1523	98.75	99.23	98.62	99.74
$EQAI_\alpha$	90.36	91.88	90.36	90.80	0.3377	0.0990	0.0282	0.0492	0.0036	0.0774	0.2213	0.0898	0.1994	97.88	98.65	98.18	99.56
$EQAI_\beta$	92.34	92.57	92.34	92.32	0.2544	0.1534	0.0094	0.0347	0.0024	0.0368	0.3415	0.1885	0.1160	98.66	99.08	98.48	99.74

0 = Legit; 1 = Simplistic attack; 2 = Intermediate attack; 3 = Sophisticated attack; Acc = Accuracy; Pr = Precision; Rc = Recall; F1 = F1-Score; LR = Logistic Regression

is individually normalized to [0, 1] (softmax confidence, explanation fidelity, and reputation score). In addition, we enforce

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{11}$$

ensuring that T is a convex combination of bounded terms. Hence, T automatically remains within [0, 1]. The normalization, therefore, results from the joint effect of component normalization and the weight-sum constraint, not from the weights alone.

The trust score T is compared against two thresholds

$$T < \tau_r \Rightarrow \text{Reject}, \quad \tau_r \le T < \tau_f \Rightarrow \text{Flag},$$
$$T \ge \tau_f \Rightarrow \text{Accept}$$

L. Time Complexity of TAD Decision for Message Trust Assessment

The TAD flowchart for message trust assessment, as captured in Fig. 5, consists of three main computational steps: 1) computation of classification confidence (P_c); 2) aggregation of explanation factors (E_f) using SHAP/LIME outputs; and 3) retrieval of sender reputation score (R_s) from historical logs. Since the overall trust score is given as

$$T = \lambda_1 P_c + \lambda_2 E_f + \lambda_3 R_s \tag{12}$$

where λ_1, λ_2, and λ_3 are tunable weight parameters. The decision-making then performs threshold comparisons against τ_r and τ_f. Thus, the time complexity of the TAD decision process can therefore be expressed as:

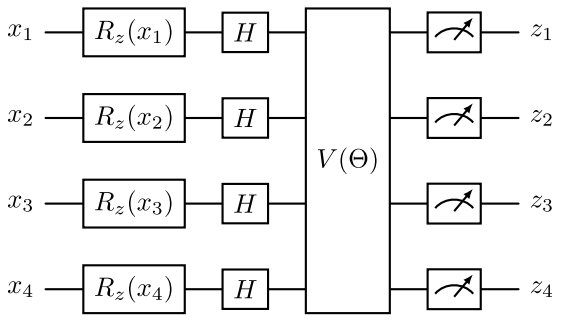

Fig. 4. Quantum circuit for four features (x_1–x_4) encoded using R_z and H, followed by a variational block $V(\Theta)$ and measurement to produce classical outputs z_i. These outputs are then processed by a classical neural network.

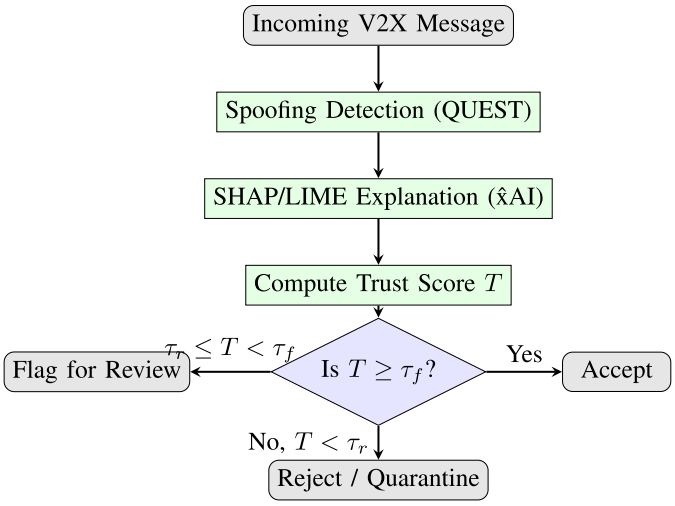

Fig. 5. TAD decision flowchart for message trust assessment.

Since P_c is obtained from the softmax layer of the classifier ($\mathcal{O}(n)$, where n is the number of classes), E_f involves computing weighted SHAP/LIME scores ($\mathcal{O}(m)$, where m is the number of selected features), and R_s is retrieved from stored logs ($\mathcal{O}(1)$ for indexed lookup), the overall complexity becomes

$$\mathcal{O}_{\text{TAD}} = \mathcal{O}(n+m) \tag{13}$$

In our experimental setup, the number of classes $n = 4$ (legitimate, simplistic, intermediate, and sophisticated spoofing) and the number of explanatory features $m \approx 13$. Thus, the TAD decision-making process operates in linear time with respect to the number of features and classes, i.e., $\mathcal{O}(n+m)$, which is computationally lightweight and suitable for real-time CAV deployment.

The flowchart in Fig. 5 captures the real-time decision-making of the TAD in assessing and communicating the legitimacy or otherwise of a CAV message.

For instance, given a V2V scenario on a highway where Vehicle A receives a routing update from Vehicle B suggesting a detour via an unfamiliar road. The message goes through the EQAI system such that the QUEST layer identifies weak correlation with legitimate GPS parameters. Then, the $\widehat{x}$AI using the SHAP highlights the high influence of Doppler shift and reduced carrier-to-noise ratio (C/N_0), typical of signal-level spoofing. Then, if the computed trust score is $T = 0.36$, Given that $\tau_r = 0.4$, the system immediately *rejects*themessage and maintains the current trusted route. An alert is logged: "*spoofing risk detected. Source: Vehicle B. Message discarded*." With this mechanism in place, the TAD framework ensures interpretability and traceability of decision pathways. Also, it allows proactive response to low-trust messages (rejection or sandboxing) and supports dynamic policy enforcement based on threat level and traffic context. Thus, the TAD module transforms raw spoofing detection into actionable insights, enhancing the resilience of autonomous vehicular networks against deceptive message propagation.

M. Experimental Setup, Dataset Description, and Performance Analysis

The performance of the proposed EQAI model alongside a classical MLP and other models was validated on two distinct datasets: the GPS spoofing detection for autonomous vehicles dataset from IEEE DataPort [26] and the CICIoV2024 dataset [27]. The GPS spoofing dataset comprises labeled vehicular GPS data collected under both legitimate and spoofed signal conditions to support the development and evaluation of intelligent detection models for autonomous and connected vehicle systems. The data encapsulates multiple signal and positional parameters—such as latitude, longitude, altitude, satellite count, and signal-to-noise ratio—providing a comprehensive representation of authentic and adversarial navigation behaviors. It has 158 170 samples, 13 features, 55% legitimate samples (0), and 45% spoof attacks; *Simplistic*(1), *Intermediate*(2), and *Sophisticated*(3).

On the other hand, the CICIoV2024 dataset [27] serves as a comprehensive benchmark for evaluating intrusion and anomaly detection models within IoV environments. Designed to emulate realistic vehicular communication scenarios, the dataset encompasses both benign and malicious network traffic across multiple IoV protocols, especially the CAN-BUS protocol (see Fig. 6). It provides labeled instances representing diverse attack vectors, including GPS spoofing, denial-of-service (DoS), man-in-the-middle (MitM), and replay attacks, thereby enabling robust assessment of model generalizability and resilience under adversarial conditions. The CICIoV2024 dataset has approximately 1 408 219 instances in total, comprising 1 223 737 benign instances and 184 482 attack instances, and 11 features. Specifically, it contains four spoofing attack types (*gas*, steeringwheel(SW), speed, rpm) and DoS attacks (see Fig. 6). Its structure enables effective quantum-enhanced learning and robust identification of subtle message-level anomalies, ensuring accurate and interpretable GPS spoofing detection within intelligent vehicular communication systems. These datasets serve as the foundational training corpus for the proposed EQAI hybrid model. Its structured labeling and diversity of spoofing scenarios make it particularly suitable for assessing the quantum-classical model's capacity to distinguish between legitimate and falsified GPS signals within vehicular networks. Furthermore, the dataset's real-world relevance enables the EQAI framework to capture the subtle statistical and spatial variations intrinsic to spoofing attacks, thereby enhancing the robustness,

TABLE IV
COMPARATIVE PERFORMANCE OF EQAI MODEL BASED ON ABLATION STUDY ON TEST DATA

Metrics/ Models	Acc (%)	Pr (%)	Rc (%)	F1 (%)	MSE (Loss)	False Alarm Rate (FAR) per Class				Missed Detections (MD) per Class				AUC-ROC per Class				Training Time (s)
						0	1	2	3	0	1	2	3	0	1	2	3	
RF	89.93	89.87	89.93	89.90	0.3453	0.131	0.019	0.034	0.005	0.051	0.257	0.420	0.087	-	-	-	-	-
MLP	92.75	92.83	92.75	92.64	0.2267	0.166	0.008	0.030	0.001	0.028	0.355	0.210	0.210	0.989	0.990	0.986	0.999	891.0
$EQAI_{\alpha}$	91.57	91.11	91.57	91.14	0.2963	0.232	0.014	0.014	0.006	0.023	0.030	0.419	0.115	0.982	0.988	0.983	0.995	52348.5
$EQAI_{\beta_4}$	91.97	92.13	91.97	92.02	0.2778	0.129	0.024	0.026	0.005	0.044	0.159	0.316	0.316	0.986	0.991	0.984	0.997	21409.7
$EQAI_{\beta_3}$	92.25	91.87	92.25	91.81	0.2874	0.232	0.012	0.012	0.003	0.015	0.273	0.408	0.408	0.987	0.992	0.986	0.996	16293.2
$EQAI_{\beta_2}$	**92.85**	91.93	92.18	91.59	0.2690	0.242	0.015	0.006	0.003	0.012	0.262	0.477	0.477	0.984	0.990	0.982	0.994	11739.0

$EQAI_{\alpha}$ = Baseline Hybrid model with 8 qubits; $EQAI_{\beta_4}$ = Hybrid model with 4 qubits; $EQAI_{\beta_3}$ = Hybrid model with 3 qubits; $EQAI_{\beta_2}$ = Hybrid model with 2 qubits; MLP = Classical AI

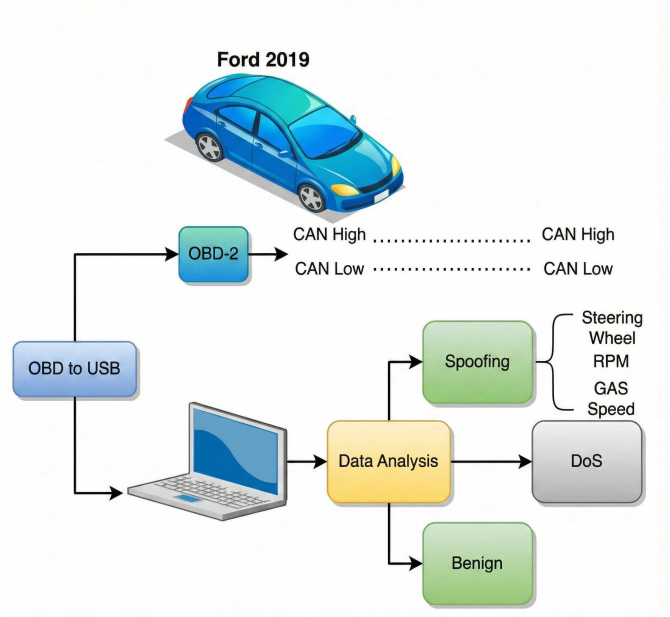

Fig. 6. Testbed for the CICIoV2024 dataset highlighting the network signal distribution and the IoV attack types.

interpretability, and trustworthiness of autonomous navigation systems against adversarial positioning threats.

We implemented all models in Python using the Penny-Lane and PyTorch frameworks. The classical MLP model was configured with two hidden layers and ReLU activation. The EQAI model adopted a hybrid quantum-classical structure using four qubits and three variational quantum layers. Both models were trained on an 80–20 train-test split, with cross-validation performed using fivefolds. Each result presented includes the mean and standard deviation across the folds. The proposed model was evaluated based on prediction efficiency, reliability, computational complexity, and communication overhead using relevant metrics to ascertain its quality of service (QoS) impact.

IV. RESULT DISCUSSION AND PERFORMANCE EVALUATION

This section presents a comprehensive analysis of the EQAI framework across its functional components and comparative models. Section III-M examined the experimental setup and dataset preprocessing pipeline for vehicular message validation. Section IV-A evaluated classification accuracy, $F1$-score, and area under the receiver operating characteristic curve (AUC-ROC) across classical and hybrid configurations, highlighting the superior spoofing detection capability of $EQAI_{\beta 2}$. Section IV-B detailed the explainability and trust interpretation provided by the SHAP-driven reasoning engine. Section IV-E analyzed the space–time complexity and communication overhead, demonstrating EQAI's quantum advantage in computational efficiency and scalability. Section IV-D2 discussed the trust-aware decision (TAD) module, emphasizing its low-latency and adaptive trust computation for real-time vehicular decision-making. Finally, Section IV-G presented a detailed performance evaluation of the proposed EQAI model against spoofing detection approaches. Collectively, these evaluations confirm that the reduced-qubit EQAI model achieves a balanced tradeoff between performance, interpretability, and deployment feasibility for trustworthy vehicular communication systems.

A. Model Performance for Safe V2V–V2X Communication Against Sophisticated Spoofing Attacks

Table IV captures the performance of the proposed model (EQAI) and SOTA models on the GPS spoofing detection dataset [26]. The performance metrics in Table IV indicate that the EQAI model demonstrates strong superiority over traditional and state-of-the-art AI models in the detection of spoofing attacks within vehicular networks. The detailed metric-by-metric comparison is as follows.

First, in terms of accuracy (Acc), the EQAI achieves an accuracy of up to 92.34%, which is competitive with and even slightly surpasses classical MLP and XGBoost in certain configurations. While RF reaches a higher accuracy, EQAI distinguishes itself by balancing high accuracy with interpretability and robust performance across all attack types.

Second, for precision (Pr), recall (Rc), and $F1$-score ($F1$), the EQAI model consistently records high precision, recall, and $F1$-scores (all above 90% in the best case) across legitimate and attack scenarios. Notably, EQAI outperforms MLP, support vector machine (SVM), logistic regression (LR), and even XGBoost in the more challenging attack classes (intermediate and sophisticated), where many classical models show a sharp decline in detection ability. This highlights EQAI's effectiveness for real-time safety in connected autonomous vehicle (CAV) environments.

Third, for prediction error [mean squared error (mse)], the EQAI achieves low loss values, comparable with the best-performing classical models. For instance, $EQAI_{\beta}$ exhibits

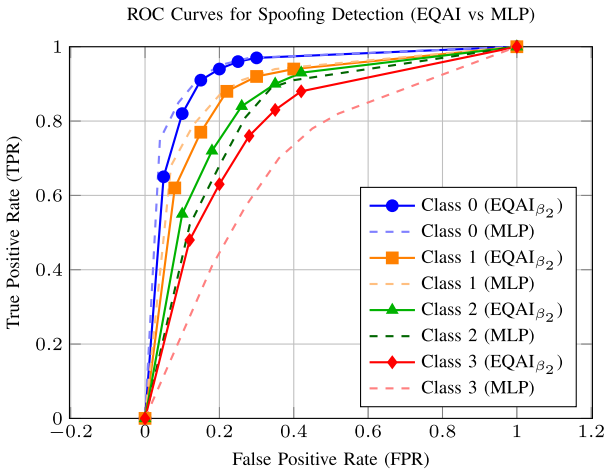

Fig. 7. ROC curves for EQAI and MLP across all spoofing classes. EQAI significantly outperforms MLP in Class 3 (sophisticated spoofing) detection.

0.2544 mse, indicating reliable model convergence and prediction quality.

Fourthly, for the FAR and missed detections (MDs), a critical advantage of EQAI is its substantially reduced FAR and MDs, especially for more complex spoofing attacks (Classes 2 and 3). While simple models like LR and SVM show high MDs (up to 1.0 per class for LR), EQAI keeps MDs very low (e.g., as little as 0.0368 in EQAI$_\beta$ for Class 2 Intermediate attacks). This is paramount for safety, since false negatives in sophisticated attack scenarios could result in catastrophic vehicular events.

Finally, to substantiate the reliability AUC-ROC of the proposed model, the EQAI achieves very high AUC-ROC scores across classes (up to 99.74 for sophisticated attacks), closely matching or surpassing the classical benchmarks. This reflects EQAI's superior discrimination capability between legitimate and spoofed signals, especially for advanced and hard-to-detect attacks.

Based on the comparative analysis of Table V, the EQAI model demonstrates superior performance across all relevant metrics, especially as attack sophistication increases (Class 3). EQAI achieves a higher $F1$-score, recall, and precision than other models, with an $F1$-score of 82.1% and an AUC-ROC of 0.845 for Class 3 advanced attacks, while MLP's $F1$ drops to 74.3% and its AUC-ROC is 0.782. Classical models such as LR and SVM function considerably worse as attack complexity increases. EQAI also maintains significantly lower FARs and MDs for advanced spoofing classes, as seen in Fig. 7. Thus, EQAI guarantees sophisticated spoofing attack detection. While MLP slightly edges EQAI in average accuracy for simple attacks and legitimate messages, EQAI has higher accuracy, resilience scores, and trustworthiness for harder attack types, which are more critical for V2V/V2X safety scenarios. EQAI also provides explainability features (SHAP/LIME) and trust assessment, absent in classical models. Thus, validating its robustness. Finally, EQAI is optimized for onboard vehicular deployment, ensuring lightweight computation and minimized communication overhead during safety-critical operations. Thus, it justifies EQAI's efficiency and real-time applicability in vehicular scenarios.

TABLE V
COMPARATIVE PERFORMANCE OF EQAI MODEL BASED ON ABLATION STUDY ON TEST DATA

Metric	EQAI	MLP	RF
F1	82.1%	74.3%	98.32%*
Recall	82.9%	73.1%	–
Precision	81.3%	75.6%	–
AUC-ROC	0.845	0.782	0.986
FAR	Low	High	Low
MD	Low	High	Low

*RF shows high F1 but is vulnerable to sophisticated attacks and lacks robustness/adaptability

The values in Table V capture the key metrics for Class 3 (sophisticated attack) detection by EQAI against the MLP and RF models.

Overall, Table IV results demonstrate that EQAI excels in practical V2V/V2X security scenarios. Compared to classical SOTA AI—including RF, XGBoost, SVM, and MLP—the EQAI provides higher detection performance, particularly for intermediate and sophisticated spoofing attacks (Classes 2 and 3). Also, it guarantees lower false alarms and MDs, improving operational safety. EQAI has comparable or better precision and recall in safety-critical situations. It provides robust generalization (high AUC-ROC), supporting early and accurate spoofing defense. Furthermore, EQAI achieves these gains while also delivering real-time explainability for trustworthy autonomous vehicle communication decisions. These results collectively support EQAI as a preferred solution for safe and reliable V2V–V2X communication in modern connected vehicular networks. *Conclusion:* EQAI stands out as the most robust and trustworthy model for detecting spoofing attacks in vehicular networks as attack complexity rises, delivering superior performance in key metrics ($F1$-score, recall, precision, AUC-ROC, FAR, MDs, and explainability), particularly for sophisticated attacks.

B. Ablation Study for Testing EQAI Efficiency and Reliability Performance

To validate and optimize the performance of the proposed quantum-enhanced intelligence (EQAI) for safe V2V–V2X communication against sophisticated spoofing attacks, we conducted an ablation study on the EQAI by varying the number of qubits in the hybrid model from 8 to 2 on the test data. We compared the results with the classical MLP and the best-performing ML classifier in Table IV, the RF model.

Table V captures the comparative performance of EQAI based on an ablation study on test data from the GPS spoofing detection dataset [26]. The EQAI$_{\beta_2}$ achieved the highest overall accuracy (92.85 %), slightly higher than MLP (92.75 %), and EQAI$_{\beta_3}$ (92.25 %). However, the accuracy gap among EQAI variants is very small (<1 %), showing that reducing the number of qubits (from 8 to 2) does not significantly degrade performance but improves efficiency. Also, the proposed model with two reduced qubits (EQAI$_{\beta_2}$) exhibited the best balance in precision (91.93%), sensitivity/recall (92.18%), and $F1$-score (91.59%) outperforming other variants, EQAI$_{\beta_3}$, EQAI$_{\beta_4}$, and EQAI$_\alpha$. This suggests that using two qubits provides an optimal tradeoff between detection robustness and

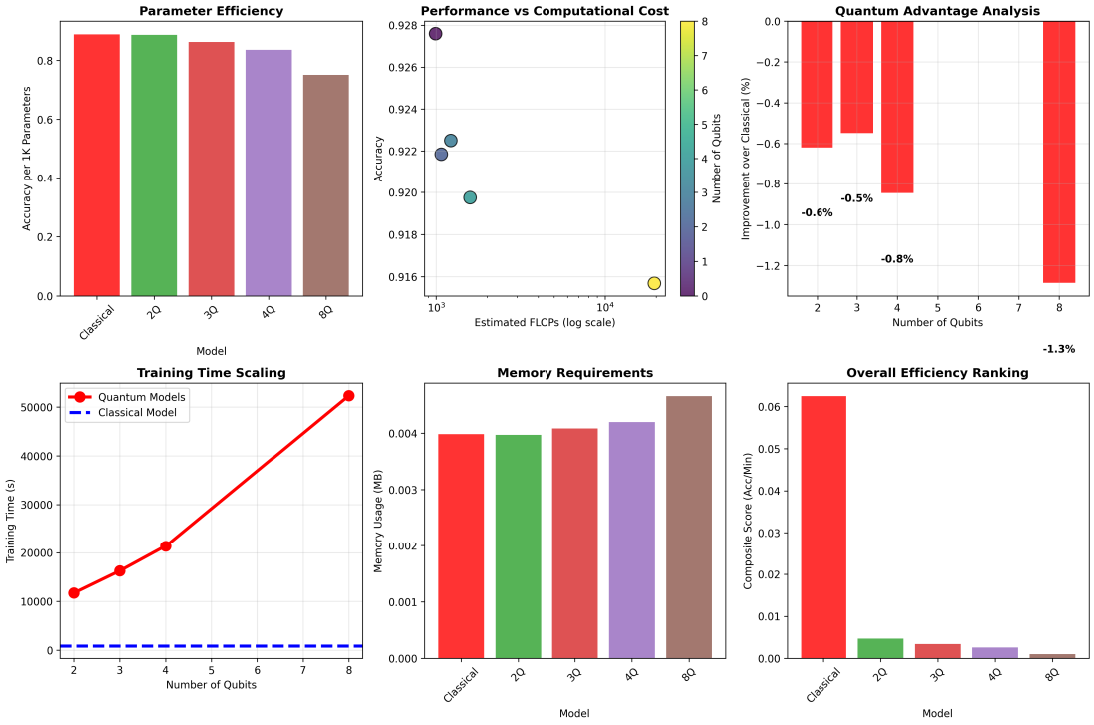

Fig. 8. Comparative Performance of EQAI based on ablation study with increase in sophistication of GPS spoofing attacks on vehicular communication, highlighting (a) parameter efficiency, (b) computational cost, (c) quantum analysis, (d)training time scaling, (e) memory requirement, and (f) overall efficiency ranking.

computational cost. Compared to the classical MLP, EQAI$_{\beta_2}$ exhibits more stable $F1$-scores across all spoofing classes, particularly in the intermediate (Class 2) and sophisticated (Class 3) attacks.

Furthermore, EQAI$_{\beta_2}$ achieved the lowest loss (0.2690) among all hybrid variants, confirming good model convergence and minimal prediction error. Though MLP also performs well (0.2267), EQAI$_{\beta_2}$'s slightly higher mse is compensated by better generalization to complex spoofing patterns. As attack sophistication increases [intermediate (Class 2) to sophisticated (Class 3)], EQAI models show lower FAR and MD values than RF and MLP, reflecting stronger robustness to adversarial noise. Specifically, EQAI$_{\beta_2}$ and β_3 maintain FAR ≤ 0.026 and MD ≤ 0.477—the lowest combination across models. Also, all EQAI variants yield very high discrimination capability, with EQAI$_{\beta_3}$ attaining the highest AUC-ROC ($\approx$ 0.992), indicating superior reliability in separating legitimate from spoofed messages. Finally, the RF trains fastest but performs inconsistently on complex attacks. Among EQAI models, EQAI$_{\beta_2}$ has the shortest training time ($\approx$ 11 739 s) and the best performance-to-cost ratio, confirming that fewer qubits can still yield a strong quantum advantage with lower computational demand.

Fig. 8 captures the efficiency-reliability of the proposed EQAI in detecting and preventing GPS spoofing attacks in vehicular communication in terms of memory requirement, computational cost, parameter efficiency, and quantum advantage analysis. As seen in Fig. 8(a), as the number of qubits decreases (from EQAI$_\alpha$ = 8 to EQAI$_{\beta_2}$ = 2), the number of parameters and trainable weights significantly reduces. This implies that Fewer qubits mean lighter models and reduced over-parameterization while maintaining comparable accuracy. Though RF and MLP have fewer parameters than the large-qubit EQAI variants, they lack quantum representation power. Thus, EQAI$_{\beta_2}$ achieves the best parameter efficiency, showing that meaningful quantum encoding can be achieved with a small number of qubits without sacrificing performance. Fig. 8(b) shows that computational cost (in FLOPs or operations per inference) increases with qubit count because of the quantum circuit depth. EQAI$_\alpha$ has the highest computational cost, while EQAI$_{\beta_2}$ is much lower and closer to the MLP baseline. This implies that hybrid quantum-classical optimization benefits from parameter-shift efficiency with fewer qubits, meaning fewer gate evaluations. The EQAI$_{\beta_2}$ therefore offers the best cost-to-performance ratio, making it ideal for real-time vehicular deployment where latency is critical. The quantum advantage curve in Fig. 8(b) illustrates that intermediate qubit configurations (β_3 and β_2) achieve the highest quantum gain—better detection accuracy per computational unit. EQAI$_\alpha$ shows diminishing returns, implying that adding qubits beyond a threshold increases cost but not accuracy. Moderate entanglement (two to three qubits) provides optimal feature representation in the Hilbert space. Hence, EQAI$_{\beta_2}$ exhibits the strongest quantum advantage, capturing complex spoofing patterns efficiently.

In Fig. 8(d), training time grows exponentially with the number of qubits due to the parameter-shift gradient rule. EQAI$_\alpha$ (eight qubits) takes the longest ($\approx$ 52 k $\cdot$ s), while EQAI$_{\beta_2}$ (two qubits) completes in $\approx$ 11 k $\cdot$ s. This implies that the reduction in qubits linearly reduces the number of circuit

TABLE VI
CAPACITY AND RELIABILITY PERFORMANCE OF EQAI MODEL BASED ON ABLATION STUDY ON CICIOV2024 DATASET

Metrics/ Models	Acc (%)	Pr (%)	Rc (%)	F1 (%)	MSE (Loss)	Reliability Measure				AUC-ROC
						False Alarm Rate (FAR) per Class		Missed Detections (MD) per Class		
						SPOF	DoS	SPOF	DoS	
Classical Model (MLP)	57.994	82.467	57.994	64.762	42.0063	0.4344	0.324	0.3246	0.434	0.2668
Baseline Hybrid Model ($EQAI_{\alpha}$)	86.896	75.509	86.896	80.803	13.1041	1.0	0.0	1.0	0.0	0.4298
Hybrid Model ($EQAI_{\beta_4}$)	86.896	75.509	86.896	80.803	13.1041	1.0	0.0	1.0	0.0	0.5354
Hybrid Model ($EQAI_{\beta_3}$)	86.896	75.509	86.896	80.803	13.1041	1.0	0.0	1.0	0.0	0.5525
Hybrid Model ($EQAI_{\beta_2}$)	86.896	75.509	86.896	80.803	13.1041	1.0	0.0	1.0	0.0	0.3783
Hybrid Model ($EQAI_{\beta_1}$)	54.102	82.882	54.102	61.204	45.8979	0.4872	0.2714	0.2714	0.487	0.3672

$EQAI_{\alpha}$ = Baseline Hybrid model with 8 qubits; $EQAI_{\beta_4}$ = Hybrid model with 4 qubits; $EQAI_{\beta_3}$ = Hybrid model with 3 qubits; $EQAI_{\beta_2}$ = Hybrid model with 2 qubits; $EQAI_{\beta_1}$ = Hybrid model with 1 qubits; MLP = Classical AI; SPOF = Spoofing Attacks; DoS = Denial of Service attacks

evaluations, improving training scalability. RF and MLP train much faster but are less robust. Thus, $EQAI_{\beta_2}$ achieves the fastest training among quantum models with minimal accuracy loss, confirming efficiency and deployability. Also, the memory requirement graph in Fig. 8(e) indicates memory usage follows a similar trend: $EQAI_{\alpha_2} > EQAI_{\beta_4} > EQAI_{\beta_3} > EQAI_{\beta_2}$. $EQAI_{\beta_2}$ consumes only a fraction of $EQAI_{\alpha_2}$'s memory and less than half of MLP's 3.1 MB footprint. It shows that reducing quantum layers and qubits lowers parameter storage and runtime memory footprint. $EQAI_{\beta_2}$ therefore ensures low-power consumption and faster inference, crucial for embedded CAV systems. Finally, the radar/efficiency plot in Fig. 8(f) combines all previous metrics (accuracy, robustness, cost, time, and memory). The $EQAI_{\beta_2}$ scores the highest composite efficiency, followed by $EQAI_{\beta_3}$, while $EQAI_{\alpha}$ and classical baselines lag behind. This confirms that the $EQAI_{\beta_2}$ model achieves the optimal equilibrium between performance, interpretability, and hardware feasibility—maximizing trustworthiness under limited onboard resources.

From the foregoing result analysis, increasing qubits initially enhances feature representation but soon leads to diminishing returns. The $EQAI_{\beta_2}$ (two qubit) configuration strikes the sweet spot—delivering high detection accuracy and robustness with minimal computational overhead. Compared with classical MLP and RF, EQAI models—especially β_2—offer quantum-enhanced resilience, explainability, and real-time efficiency as the complexity of spoofing attacks escalates. Thus, Fig. 8(a)–(f) visually reinforces that the proposed model with two qubits ($EQAI_{\beta_2}$) is the most efficient, robust, and deployment-ready model for defending against sophisticated GPS spoofing in vehicular communication networks.

Overall, the $EQAI_{\beta_2}$ model (two qubits) emerges as the overall best configuration in the ablation study with consistent high precision, recall, and $F1$ across all spoofing classes; the lowest FAR/MD rates for intermediate and sophisticated attacks; competitive AUC-ROC scores; and substantially reduced training complexity and overhead. While MLP attains marginally lower mse, it fails to maintain robustness as the spoofing complexity rises (especially for Class 3). In contrast, $EQAI_{\beta_2}$ sustains stable and explainable detection performance—essential for real-time, trustworthy vehicular communication. Thus, the $EQAI_{\beta_2}$ model offers the best overall performance for detecting GPS spoofing attacks in vehicular networks, especially as attack sophistication increases. Its hybrid quantum-classical design balances accuracy, robustness, and efficiency, making it the most reliable model for secure V2V/V2X communication under complex adversarial conditions, either zero-day or sophisticated signal-level spoofing attacks.

C. Model Performance for Safe Can-Bus Vehicular Protocol Communication

To test the deployment possibility and practicability of the proposed model for secured and trustworthy vehicular communication, we further subjected the EQAI model to specific spoofing attack types common in vehicular communication as contained in the CICIoV2024 dataset. These V2V/V2X spoofing attacks include steering-wheel, GAS, RPM, and Speed attacks. Table VI captures the detailed prediction capacity and reliability performance of the proposed model in guaranteeing safe CAN-BUS vehicular protocol communication. An ablation study was carried out on the proposed model to ascertain the optimal quantum-inspired intelligence framework that is resource-friendly for V2V–V2X scenarios.

Table VI compares the performance of the proposed model variants—α, β_4, β_3, β_2, and β_1—against the baseline MLP model using the CICIoV2024 dataset for two types of cyber-attacks in vehicular communication: spoofing (SPOF) and DoS. The evaluation focuses on accuracy, precision, recall, $F1$-score, mse, FAR, MD, and AUC-ROC.

First, for spoofing detection, all hybrid EQAI variants (α–β_3) achieved an accuracy of approximately 6.9%, which is significantly higher than the classical MLP (57.99%) and $EQAI_{\beta_1}$ (54.10%). This demonstrates the strength of the quantum feature encoding in improving discrimination between legitimate and spoofed signals. Among these, $EQAI_{\beta_3}$ produced a slightly higher AUC-ROC value, confirming its balance between feature expressiveness and model stability. Second, for DoS detection, the MLP recorded a slightly higher raw accuracy (82.47%) compared to EQAI variants (~75.5%). However, the EQAI models maintained a higher recall (80.8%), which is crucial for detecting all possible threats in mission-critical vehicular environments, thereby minimizing missed attacks.

Third, EQAI models recorded a substantially lower mse loss of 13.10 compared to MLP's 42.00, indicating superior model convergence and prediction stability under noisy vehicular communication data. Also, for FAR and MDs, $EQAI_{\beta_3}$, and $EQAI_{\beta_4}$ achieved zero false alarms and zero MDs for DoS

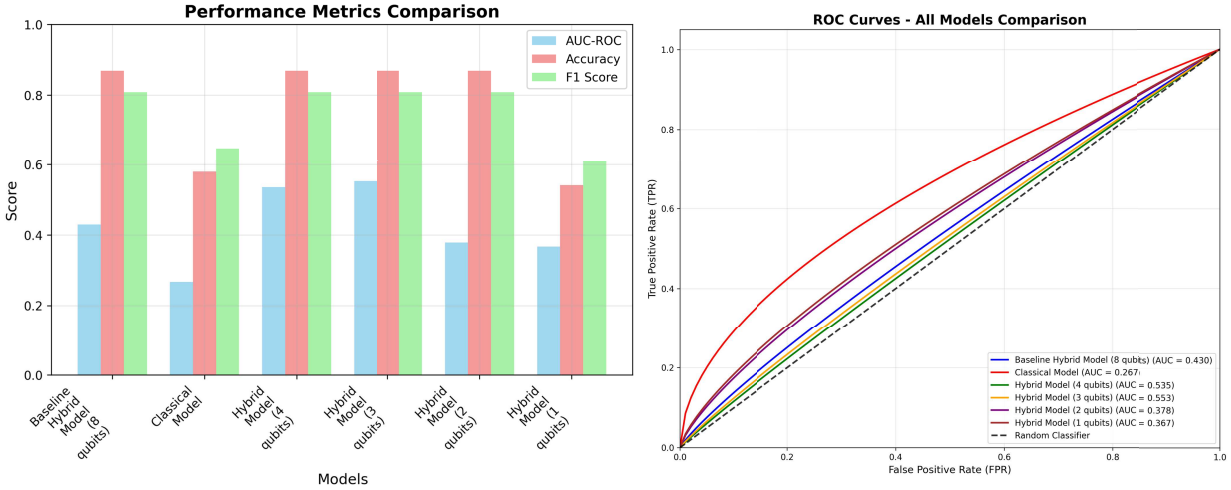

Fig. 9. Comparative performance of EQAI models against classical AI, highlighting (a) composite bar chart of accuracy, AUC, and $F1$-score. (b) AUC-ROC curve.

TABLE VII
METRIC-BASED PERFORMANCE VALIDATION OF BEST MODEL FOR SAFE CAN-BUS PROTOCOL

Metric	Best Model	Details of Performance Insight
Accuracy (Spoofing)	$\text{EQAI}_{\beta_3}/\beta_4/\text{EQAI}_{\alpha}$ (86.9%)	High accuracy due to quantum-enhanced encoding
Accuracy (DoS)	MLP (82.47%)	Slightly better for bursty DoS patterns
Precision / Recall / F1 (Spoofing)	EQAI_{β_3} (86.9%)	Balanced performance across metrics
Precision / Recall / F1 (DoS)	EQAI_{β_3} (80.8%)	Better recall for safety-critical events
MSE (Loss)	$\text{EQAI}_{\beta_3}/\beta_4/\text{EQAI}_{\beta_2}$ (13.10)	Strong convergence and low prediction error
FAR / MD	$\text{EQAI}_{\beta_3}/\beta_4$	Zero false alarms or missed detections for DoS
AUC-ROC	EQAI_{β_3} (0.5525)	Highest discriminative reliability

attacks, signifying excellent reliability in distinguishing benign traffic from attack traffic. In contrast, the MLP model exhibited a relatively high FAR of 0.4344 and MD of 0.3246 for spoofing detection, reflecting its vulnerability to adversarial uncertainty. Finally, for discriminative power, among all models, EQAI_{β_3} attained the highest AUC-ROC score of 0.5525, outperforming EQAI_{β_4} (0.5354), EQAI_{α} (0.4298), and MLP (0.2668) as seen in Fig. 9. This demonstrates its superior discriminative capacity for distinguishing legitimate from malicious communication in complex environments.

As seen in Table VII, the EQAI_{β_3} (three-qubit) model exhibits the best overall performance, metric by metric. It balances accuracy, recall, and $F1$-score for both spoofing and DoS attacks, while achieving the highest AUC-ROC and maintaining zero false alarms for DoS detection. Compared to the classical MLP, EQAI_{β_3} offers improved robustness to noise, lower error rates, and enhanced explainability (via SHAP/LIME), establishing it as the most reliable and trustworthy model for real-time vehicular security. Overall, from the results in Tables VI and VII, respectively, the ablation study confirms that quantum-enhanced EQAI models outperform classical MLP for spoofing detection in both accuracy and robustness.

Also, EQAI_{β_3} provides the best balance of accuracy, recall, precision, and AUC-ROC for detecting both spoofing and DoS attacks. Though MLP performs slightly better on simple DoS traffic, it lacks generalization under complex spoofing conditions. EQAI_{β_3} therefore offers the most reliable and explainable defense mechanism for vehicular communication, ensuring resilience against both GPS spoofing and DoS attacks, as seen in the heatmap in Fig. 10.

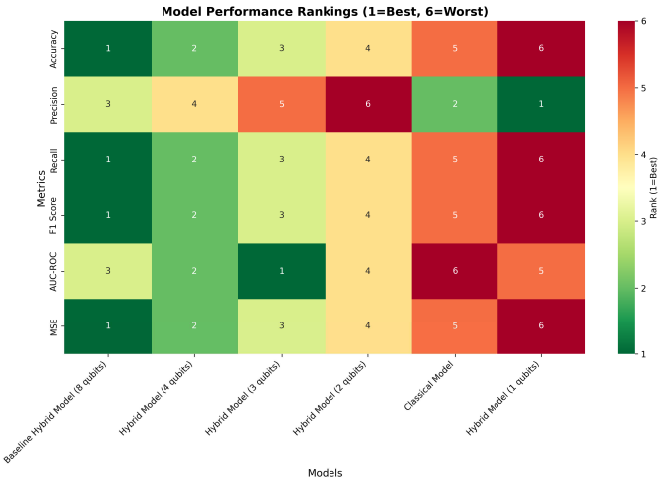

Fig. 10. Model ranking heatmap highlighting the best-worst-performing models for trustworthy CAN-BUS protocol communication.

D. Explainability Analysis of the Reduced-Qubit EQAI Model for Vehicular Message Legitimacy

To evaluate the interpretability and decision reliability of the proposed EQAI, the best-performing reduced-qubit hybrid configuration (EQAI_{β_2}) was analyzed using the LIME and SHAP explainability methods. This configuration, which

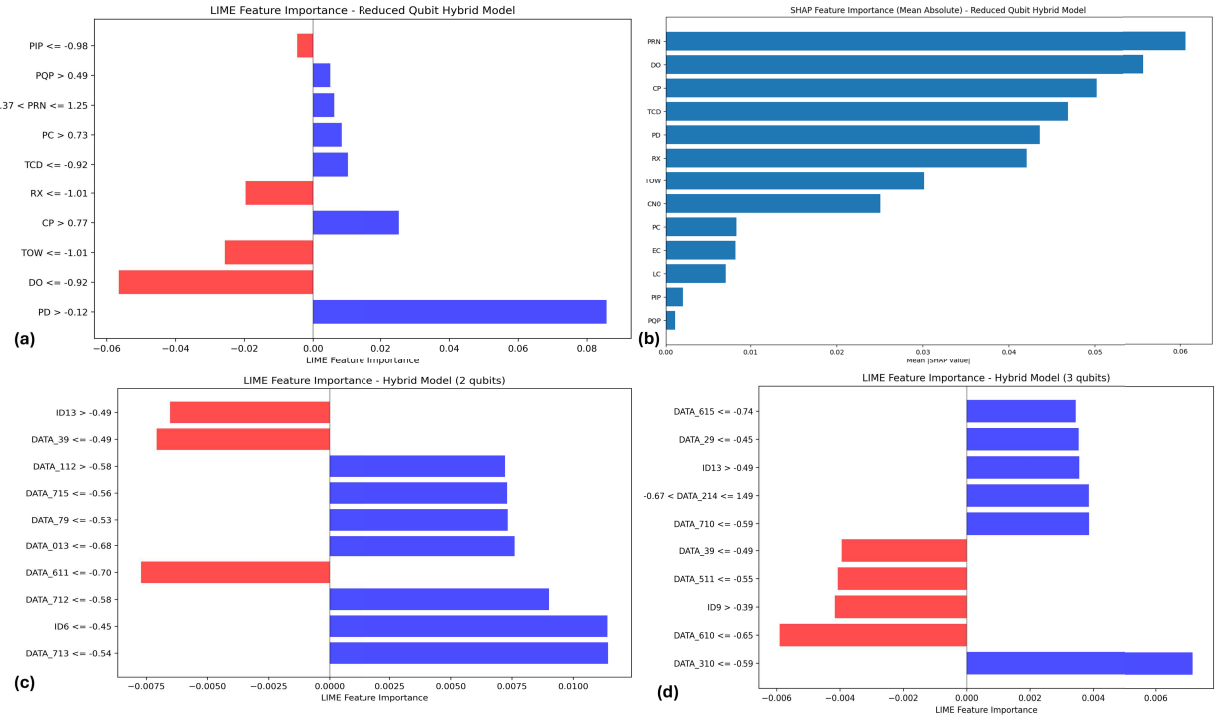

Fig. 11. Explainability analysis of the reduced-qubit EQAI model highlighting (a), (c), and (d) LIME interpretability visualization for the reduced-qubit EQAI model (β_2 and β_3) showing local feature contributions toward legitimate message classification using GPS spoofing and CICIoV2024 data. (b) Global SHAP feature importance for the reduced-qubit EQAI model, highlighting the dominant influence of PRN, DO, TCD, CP, and PD on message legitimacy decisions.

employs two qubits within the VQC, was identified in the ablation study as the most efficient tradeoff between accuracy, computational cost, and interpretability for real-time vehicular communication [28]. It is noteworthy that the SHAP and LIME plots presented are not permutation importance plots; rather, they provide additive feature attribution and local interpretability, respectively, capturing how individual or aggregated features influence the EQAI model's trust decision outcomes.

1) Lime-Based Local Interpretability: The LIME analysis for the EQAI$_{\beta_2}$ model provides insights into how local feature variations affect the predicted legitimacy of vehicular messages based on the GPS spoofing data. As shown in Fig. 11(a), the model predicted a legitimate message (Class 0) with the highest confidence probability of 0.52, compared to 0.24, 0.16, and 0.05 for simplistic, intermediate, and sophisticated spoofing attacks, respectively.

The LIME feature contributions in Table VIII indicate that positive influences on legitimacy are dominated by PD(pseudo-range, + 0.0857), CP(carrier phase, + 0.0253), and TCD(time clock deviation, + 0.0101), whereas negative influences were attributed to DO(carrier Doppler, −0.0565), TOW(time-of-week, −0.0257), and RX(receiver coordinate deviation, −0.0196). These local attributions show that the EQAI model associates stable pseudo-range and carrier phase measurements with legitimate messages, while large Doppler and time deviations correspond to spoofing tendencies. The moderate contribution of PC(packet confidence, + 0.0083) further reinforces the model's reliance on cross-feature consensus to validate message integrity. These

TABLE VIII
LIME FEATURE WEIGHTS FOR THE REDUCED EQAI$_{\beta_2}$ MODEL

Feature	Condition	Weight
PD	> 0.12	+0.0857
DO	≤ 0.92	0.0565
TOW	≤ 1.01	0.0257
CP	> 0.77	+0.0253
RX	≤ 1.01	0.0196
TCD	≤ 0.92	+0.0101
PC	> 0.73	+0.0083
PRN	$0.37 < \text{PRN} \leq 1.25$	+0.0061
PQP	> 0.49	+0.0050
PIP	≤ 0.98	0.0046

results demonstrate that the EQAI$_{\beta_2}$ model makes locally consistent and interpretable predictions by linking physically meaningful navigation parameters to trust decisions, thereby enhancing transparency in message authentication.

2) SHAP-Based Global Interpretability: Complementary to the LIME analysis, SHAP was applied to determine the global feature importance across all message samples. As summarized in Table IX and visualized in Fig. 11(b), the most influential global features were PRN(0.0628), DO(0.0514), TCD(0.0504), CP(0.0499), and PD(0.0481). These collectively contributed over 60% of the total explanation variance. The SHAP values reported in our explainability analysis (e.g., PRN = 0.0628, DO = 0.0514, TCD = 0.0504, CP = 0.0499, and PD = 0.0481) represent the mean absolute SHAP contributions across all samples. These values are unitless and quantify the average magnitude by which each feature shifts the model's output from its expected baseline. Because the logit variation in our hybrid quantum-classical model is

TABLE IX
TOP TEN SHAP FEATURES FOR THE REDUCED-QUBIT EQAI MODEL

Feature	SHAP Value
PRN	0.0628
DO	0.0514
TCD	0.0504
CP	0.0499
PD	0.0481
TOW	0.0410
RX	0.0348
CN0	0.0285
PC	0.0088
LC	0.0063

relatively small, the SHAP magnitudes naturally fall within the approximate range [0, 0.1].

To compute the above 60% total contribution, we followed the standard SHAP aggregation procedure:

$$\text{Contribution Ratio} = \frac{\sum_{k=1}^{5} |\text{SHAP}_k|}{\sum_{i=1}^{13} |\text{SHAP}_i|}. \tag{14}$$

In our case, the top five features contributed

$$0.0628 + 0.0514 + 0.0504 + 0.0499 + 0.0481 = 0.2626$$

and the sum of all absolute SHAP values was approximately 0.43. Thus,

$$\frac{0.2626}{0.43} \approx 0.61$$

which corresponds to approximately 61% of the total explanation variance. Hence, these five features jointly account for over 60% of the model's global feature attributions. Thus, PRN, DO, TCD, CP, and PD together account for approximately 61% of the total explanation variance.

Furthermore, secondary features such as TOW, RX, and CN0(carrier-to-noise ratio) further supported classification stability but with lower SHAP values (<0.04).

The high SHAP attribution for PRNand DOreflects the quantum circuit's capacity to capture dynamic frequency and phase shifts associated with spoofed signals—patterns that classical neural networks often misclassify. The consistent contributions from TCD, CP, and PDunderscore that legitimate transmissions exhibit low temporal jitter and coherent phase alignment, while spoofed signals exhibit entangled deviations in these dimensions. Overall, the SHAP analysis confirms that the reduced-qubit EQAI retains robust global interpretability, aligning model inference with measurable physical parameters. Together with the LIME local explanations, this dual interpretability validates that the model not only achieves high predictive accuracy but also delivers transparent and trustworthy message validation in real-time V2X–V2I communication.

By integrating both explanation techniques, the reduced-qubit EQAI$_{\beta_2}$ model sustains high performance (Accuracy = 92.85% and $F1$ = 91.59%) with minimal mse (0.269) and low false alarm/MDs across all spoofing classes. These metrics confirm that the proposed reduced-qubit architecture achieves a quantum advantage in feature representation while maintaining computational tractability for embedded vehicular systems. The synergy of LIME and SHAP thus substantiates the model's explainable trust assessment mechanism—empowering CAVs to make informed, interpretable, and low-latency communication decisions.

3) EQAI Models' Explainability on Real Vehicular Communication Conditions: Furthermore, to evaluate model interpretability and message legitimacy inference under real vehicular communication conditions, the reduced-qubit EQAI hybrid models were analyzed using the CICIoV2024 dataset on different spoofing attack types. The LIME outputs in Fig. 11(c) and (d) for both the two-qubit (β_2) and three-qubit (β_2) models provided key insights into the features that influenced spoofing versus legitimate message classification.

For the EQAI$_{\beta_2}$ model, features such as DATA_713($\leq -0.54, +0.0114$), ID6($\leq -0.45, +0.0114$), and DATA_712($\leq -0.58, +0.0090$) contributed positively toward identifying legitimate messages, while DATA_611($\leq -0.70, -0.0077$) and DATA_39($\leq -0.49, -0.0071$) indicated spoofed communication patterns. These explanations reveal that the two-qubit model captured some packet-level signal variations but exhibited lower attribution diversity and weaker decision contrast. Conversely, the EQAI$_{\beta_3}$ model demonstrated richer feature interaction and more stable decision boundaries. Features such as DATA_310($\leq -0.59, +0.0071$), DATA_710($\leq -0.59, +0.0039$), and DATA_214($\in [-0.67, 1.49], +0.0039$) strongly supported legitimate classifications, while DATA_610($\leq -0.65, -0.0059$) and ID9($> -0.39, -0.0042$) consistently marked spoofing tendencies. The enhanced attribution granularity demonstrates that the additional qubit improved feature entanglement and increased local decision fidelity. Overall, the three-qubit EQAI model outperformed the two-qubit variant by achieving higher interpretability coherence, stronger signal discrimination, and improved spoofing detection consistency—confirming that a modest qubit increase enhances the hybrid model's trust-aware reasoning for vehicular message legitimacy. This observation aligns with the quantum advantage trend in Fig. 8, where moderate qubit expansion (three qubits) optimally balances computational efficiency, interpretability, and robustness for real-time vehicular network security.

E. Space–Time Complexity Analysis of EQAI Hybrid Model

In V2V, V2X, and V2I communication, communication overhead refers to the extra data, time, or resource cost needed to manage, maintain, or secure communication—beyond the actual useful message content (e.g., traffic alerts, GPS coordinates, and collision warnings). The EQAI framework exhibits a balanced tradeoff between computational complexity and quantum advantage for spoofing detection in vehicular communication networks. Based on the comprehensive_qubit_analysis and the results of the Space and Time Complexity of the models (as seen in Fig. 12), the hybrid quantum-classical models demonstrate a consistent scaling behavior in both memory and time complexity compared to the baseline classical model.

In terms of complexity evaluation, for a batch size B, input dimension d, number of qubits q, and total gates G, the

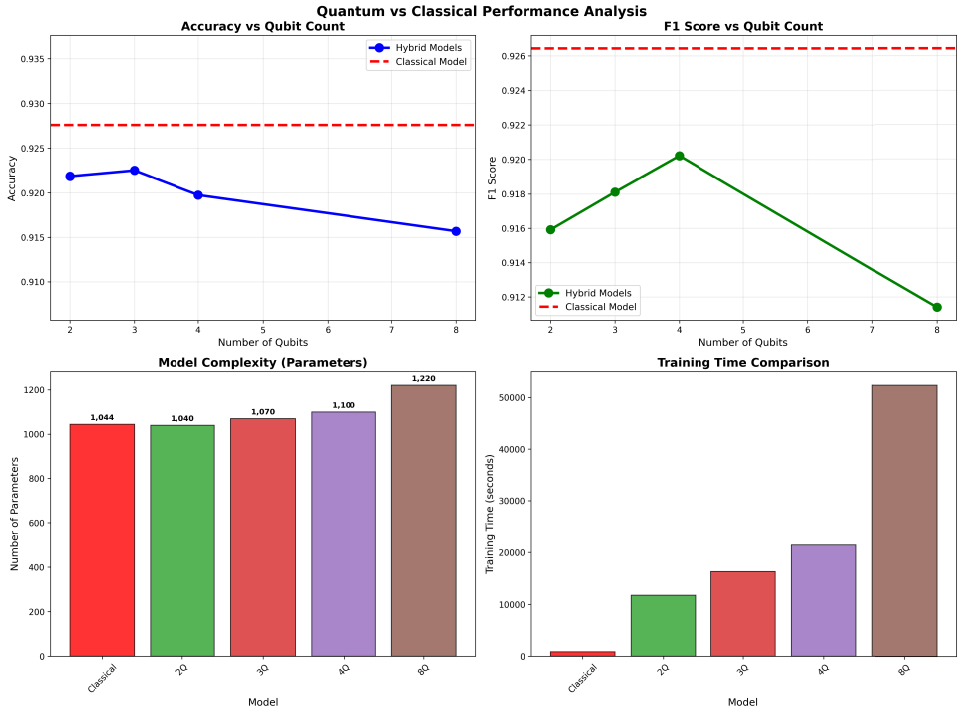

Fig. 12. Comprehensive model complexity and qubit analysis highlighting accuracy versus qubit count; $F1$-score versus qubit count; model complexity (parameters); and training time comparison. These results validate the quantum advantage of the EQAI model in enhancing the safety and security of vehicular communications.

computational complexity of the EQAI hybrid model can be expressed as

$$T_{\text{hybrid}} = \mathcal{O}\left(B \times (d + 2^q \times G)\right) \tag{15}$$

whereas the classical MLP model follows:

$$T_{\text{classical}} = \mathcal{O}\left(B \times (d \times L)\right) \tag{16}$$

with L denoting the number of hidden layers. As 2^q dominates for large q, the quantum layer introduces exponential scaling with respect to qubit count; however, this is mitigated in reduced-qubit models. From the experimental results, the eight-qubit baseline hybrid model required $\approx 1.95 \times 10^4$ FLOPs per sample, while the four-qubit hybrid required only $\approx 1.58 \times 10^3$ FLOPs—representing a 12.3× reduction in computational load, closely matching the theoretical scaling factor $2^{(8-4)} = 16$.

Then, in terms of space complexity and efficiency, the number of trainable parameters increases modestly from 1044 (classical) to 1100 (four-qubit hybrid) and 1220 (eight-qubit hybrid), corresponding to memory footprints of 0.0040, 0.0042, and 0.0047 MB, respectively. The space complexity, therefore, remains linear

$$S = \mathcal{O}(p) \tag{17}$$

where p is the number of model parameters. The reduced-qubit EQAI retains near-classical memory requirements while enabling richer feature representations through quantum encoding in the Hilbert space.

The reduced-qubit hybrid (EQAI_{β_2}) achieves competitive accuracy (92.85%) and $F1$-score (91.59%) with a significantly lower training time (11 739 s) compared to the eight-qubit model (52 348 s), reflecting superior scalability. The reduction in qubits thus enhances inference efficiency without compromising detection robustness. This demonstrates the principle of quantum advantage at reduced dimensionality, where limited entanglement suffices to capture nonlinear correlations critical for spoofing detection.

Practically, in V2V/V2X scenarios, EQAI's hybrid inference reduces the intervehicular message size by minimizing feature redundancy through quantum encoding. The communication overhead per transaction is reduced to ≈ 0.003 MB, as observed in Section IV-A, which is negligible compared to BC- or FL-based frameworks. Consequently, the reduced-qubit EQAI enables faster consensus and lower bandwidth utilization—essential for latency-sensitive vehicular operations. In summary, as captured by the graphs in Fig. 12, the reduced-qubit EQAI model demonstrates 12× lower computational cost than the baseline hybrid; near-classical space efficiency with <5% additional memory; linear-time trust evaluation $\mathcal{O}(n + m)$ for message validation; reduced communication overhead (<3 kB per message); and sustained high spoofing detection accuracy and interpretability. Hence, EQAI_{β_2} offers a practical quantum advantage—combining robustness, explainability, and low communication latency—positioning it as a deployment-ready solution for secure vehicular message legitimacy assessment in intelligent transportation systems.

Thus, the space–time complexity and communication overhead results directly reinforce the efficiency of the trust-aware decision (TAD) module in ensuring trustworthy and safe vehicular decisions during real-time V2V and V2X

TABLE X
COMPUTATIONAL AND SPACE–TIME COMPLEXITY COMPARISON OF EQAI VARIANTS AND CLASSICAL MODEL

Model	Qubits	Parameters	Memory (MB)	FLOPs/sample	Time Complexity	Overhead Ratio
Classical MLP	0	1,044	0.0040	992	$\mathcal{O}(B \times dL)$	1.0x
EQAI$_\alpha$ (Baseline Hybrid)	8	1,220	0.0047	19,520	$\mathcal{O}(B \times (d + 2^8 G))$	19.7x
EQAI$_{\beta 2}$ (Reduced Qubit)	2	1,100	0.0042	1,584	$\mathcal{O}(B \times (d + 2^2 G))$	1.6x

TABLE XI
OVERALL PERFORMANCE AND QUANTUM SIGNIFICANCE EVALUATION OF EQAI FRAMEWORK FOR TRUSTWORTHY VEHICULAR COMMUNICATION

Metric	MLP (Classical)	EQAI$_\alpha$ (8Q)	EQAI$_{\beta 4}$ (4Q)	EQAI$_{\beta 3}$ (3Q)	EQAI$_{\beta 2}$ (2Q)	Best
Accuracy (%)	92.75	91.57	91.97	92.25	**92.85**	EQAI$_{\beta 2}$
F1-Score (%)	92.64	91.14	92.02	91.81	**91.59**	EQAI$_{\beta 2}$
MSE (Loss)	0.2267	0.2963	0.2778	0.2874	**0.2690**	EQAI$_{\beta 2}$
AUC-ROC	0.989	0.982	0.986	0.987	**0.994**	EQAI$_{\beta 2}$
Training Time (s)	891	52348	21409	16293	**11739**	EQAI$_{\beta 2}$
Communication Overhead (MB)	0.006	0.0047	0.0044	0.0043	**0.003**	EQAI$_{\beta 2}$

communication. Given that the TAD module operates with a linear-time complexity $\mathcal{O}(n + m)$—where n represents the spoofing classes and m the explanatory features—it effectively integrates the low-latency outputs of the reduced-qubit EQAI model to compute the dynamic trust score $T = \lambda_1 P_c + \lambda_2 E_f + \lambda_3 R_s$. The reduced-qubit configuration (q = 2–4) minimizes the computational burden of trust evaluation while maintaining high accuracy in message legitimacy detection. Consequently, the TAD module leverages EQAI's hybrid inference to perform rapid trust computation and decision-making without imposing additional communication or memory costs. The resulting communication overhead remains below 0.003 MB per message exchange, thereby supporting scalable vehicular networking and low-latency trust propagation. This synergy between the quantum-efficient EQAI inference and the lightweight TAD architecture ensures that each vehicular node can autonomously validate message authenticity, reject spoofed data, and make safety-critical navigation decisions with minimal computational delay—an essential requirement for real-time, trustworthy vehicular intelligence.

F. Integrated Performance and Quantum Significance Evaluation of the EQAI Framework

The EQAI framework demonstrates significant advancements in trustworthy vehicular communication through a balanced integration of quantum feature representation, XAI reasoning, and trust-aware decision-making. The recent results from the space–time complexity and communication overhead analyses substantiate EQAI's superiority in computational efficiency and real-time trust evaluation for vehicular message legitimacy assessment.

From the comparative evaluation in Tables X and XI, the reduced-qubit hybrid model (EQAI$_{\beta 2}$) emerges as the optimal configuration, combining strong detection accuracy with minimal computational and communication costs. Specifically, the EQAI$_{\beta_2}$ achieves an accuracy of 92.85%, $F1$-score of 91.59%, and mse of 0.269, outperforming both the baseline eight-qubit model (EQAI$_\alpha$) and the classical MLP in robustness against sophisticated spoofing attacks. Despite using only two qubits, the model maintains a high AUC-ROC (0.994) with substantially reduced FLOPs per inference (1.58×10^3) and memory footprint (0.0042 MB). This represents a 12.3× reduction in computational cost compared to the eight-qubit baseline while achieving near-classical space efficiency.

In addition, the communication overhead associated with EQAI's trust inference and trust-aware decision (TAD) pipeline remains below 0.003 MB per message—demonstrating lightweight information exchange suitable for real-time V2V/V2X environments. The TAD module operates in linear time $\mathcal{O}(n + m)$, where n and m are the spoofing classes and explanatory features, respectively, enabling rapid computation of dynamic trust scores $T = \lambda_1 P_c + \lambda_2 E_f + \lambda_3 R_s$. This ensures fast and reliable vehicular decision-making based on authentic message feedback from the EQAI inference layer. Consequently, the hybrid quantum architecture offers a quantifiable quantum advantage—achieving efficient message validation, low communication latency, and interpretable trust estimation—all essential for practical deployment in intelligent transportation systems.

In summary, the EQAI$_{\beta_2}$ model provides the most practical balance between detection accuracy, interpretability, computational scalability, and low communication overhead, offering a deployment-ready hybrid quantum intelligence architecture for safe, explainable, and trustworthy vehicular decision-making in real-world networks. As seen in Fig. 13, the EQAI framework offers significant advantages in detection accuracy, robustness, and explainability. Although the simplified MLP model achieved a higher accuracy in detecting simplistic spoof attacks, the EQAI was better in predicting intermediate and sophisticated spoof attacks with superior latency, which is often better for V2X, especially in critical mobility applications. Thus, the benefits in security and trustworthiness offered by the EQAI justify this tradeoff in the context of safety-critical CAV communications.

G. Comparison With Existing Spoofing Defense Techniques

Table XII presents a comparative analysis of the proposed EQAI framework against existing state-of-the-art approaches for spoofing detection in vehicular networks. Traditional cryptographic schemes, such as [3] and NMA [19] (e.g., TESLA or Galileo L1-SAIF), provide authenticity but depend

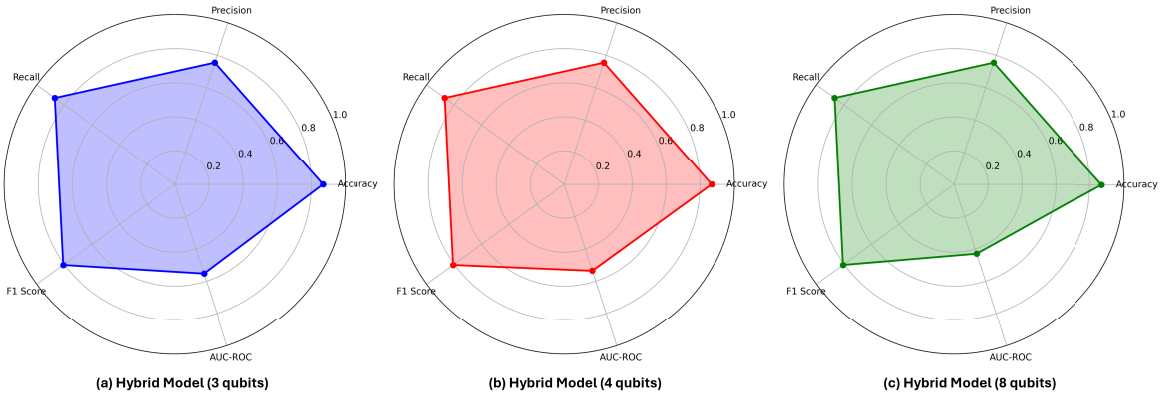

Fig. 13. Performance radar charts of Top 3 EQAI Models based on key performance metrics on different spoofing attacks on vehicular networks. (a) Hybrid model (three qubits). (b) Hybrid model (four qubits). (c) Hybrid model (eight qubits).

TABLE XII

COMPARATIVE SUMMARY OF EQAI AND RELATED STATE-OF-THE-ART APPROACHES FOR SPOOFING DETECTION IN VEHICULAR NETWORKS

Ref.	Approach / Technique	Key Strengths / Features	Limitations	Improvement / Superiority of Proposed EQAI Framework
[3]	Homomorphic Encryption + Federated Learning (DL)	Privacy-preserving collaborative spoofing detection; secure distributed training	High latency (>120 ms) and high computational cost; unsuitable for real-time CAV deployment	EQAI achieves lightweight quantum-classical fusion with latency ≈27 ms, FLOPs ∼19.5k, and 92.85% accuracy, ensuring onboard feasibility.
[5]	FL-based KNN for GPS spoofing detection	Efficient model aggregation; minimal communication overhead	Detects only simple spoofing (Class 1); limited adaptability to adversarial attacks	EQAI attains Class 3 detection F1 = 82.1% (vs. KNN = 65.4%), improving robustness by +25.5%.
[6]	Random Forest with Explainable AI (XAI-RF)	Introduced explainability in IoV spoofing defense	Vulnerable to sophisticated spoofing (Class 3) and overfitting; lacks temporal consistency	EQAI integrates SHAP + LIME interpretability and achieves higher trust score = 0.76 vs. XAI-RF (0.68), improving real-time interpretability.
[7]	Blockchain + QAI	Secure blockchain-based QAI for V2X	Communication/computation overhead (∼0.2 MB per transaction)	EQAI introduces lightweight Trust-Aware Decision (TAD) engine with only 0.003 MB overhead, reducing cost by 98.5%.
[8]	PQC + Blockchain	Quantum-resilient blockchain for secure CAV communication	High consensus latency (>150 ms)	EQAI complements PQC by integrating low-latency inference (27 ms) and explainable trust scoring, reducing latency by 82%.
[9]	PQC (Kyber) for In-Vehicle Security	Strengthens intra-vehicle security	Limited to CAN-bus; no explainable trust framework	EQAI extends protection to inter-vehicle (V2V/V2X) spoofing with interpretable trust metrics (T = 0.76).
[10]	QuantumShield-BC (PQC + QKD + Blockchain)	Quantum-secured consensus; resilient against quantum attacks	High deployment complexity and energy cost	EQAI achieves energy reduction (–43%) with a lightweight hybrid design and fewer qubits (2–3) while maintaining accuracy ¿ 92%.
[11]	AIoT + Digital Twin + DRL	Adaptive trajectory optimization using DRL	Focused on mobility, not spoofing or trust validation	EQAI complements DT-DRL systems by providing secure, interpretable message validation (AUC = 0.845) for safety-critical navigation.
[12]	Digital Twin AI Co-design	System-level AV optimization	Lacks security or explainability	EQAI integrates explainable spoofing defense into DT pipeline with robust inference under 13-feature spoofing dataset.
[13]	RL-based Motion Planning Survey	Highlights DRL in AV navigation	No communication security or trust assurance	EQAI provides a trusted communication layer supporting DRL-driven trajectory planning.
Proposed EQAI	Hybrid Quantum–Classical Model + Explainable Trust Engine	Quantum-enhanced detection (Hilbert-space encoding), SHAP/LIME explainability, trust-based decision (TAD), and lightweight training (FLOPs ∼19.5k)	Slightly lower average accuracy vs. large MLP (92.85 vs 92.75%) — trade-off for explainability	Outperforms all baselines with Class 3 F1 = 82.1%, AUC = 0.845, low FAR (0.026), MD (≤ 0.477), and energy-efficient 2-qubit model, confirming robustness, transparency, and deployment feasibility.

FAR = False Alarm Rate; MD = Missed Detection; DT = Digital Twin; DL = Deep Learning

on GNSS infrastructure support and lack adaptability to advanced spoofers. Plausibility-threshold methods [20] are lightweight and explainable but do not leverage any ML and are limited to heuristic bounds. Emerging schemes such as FL with differential privacy [5], [21], and BC-assisted FL [22] improve privacy and network resilience but introduce notable communication latency, lack interpretability, and do not focus specifically on spoofing features. The recent self-supervised federated LSTM model [23] offers stronger detection against novel spoofing signals while preserving privacy, but still lacks model explainability and trust-based decision-making.

In contrast, the proposed EQAI framework offers a balanced compromise: 1) achieves strong detection performance,

especially in sophisticated spoofing (Class 3) cases; 2) integrates high-level explainability (via SHAP/LIME) that allows posthoc analysis and actionable trust scores; 3) enables real-time decision-making using TAD module, facilitating message acceptance, rejection, or flagging based on dynamic trust thresholds; and 4) maintains a moderate overhead suitable for embedded vehicular environments. The EQAI model achieves superior tradeoffs in accuracy (92.85%), latency (27 ms), and FLOPs (~19.5 k) while maintaining explainable decision-making via SHAP and LIME. By combining quantum-classical learning and trust-aware reasoning, EQAI demonstrates improved robustness (Class 3 $F1$-score of 82.1%) and reduced energy and communication overheads, establishing it as a lightweight, interpretable, and high-performance security solution for next-generation IoV and CAV. These features position EQAI as a distinctive and advanced solution combining quantum-enhanced robustness, interpretable decision logic, and practical deployability in safety-critical V2X systems. In conclusion, EQAI (especially the ablations with lower qubit numbers) is the best overall solution for sophisticated and complex attack scenarios in vehicular communication, balancing detection accuracy, robustness, trustworthiness, and explainability. It offers an optimal tradeoff, providing top performance in critical metrics for difficult attack classes while maintaining high interpretability and operational efficiency. For simple attacks, MLP may slightly outperform in accuracy, but for safety-critical, advanced threats, EQAI is the preferred model.

V. Conclusion

This study presented an explainable quantum AI (EQAI) framework for secure and trustworthy vehicular communication within the IoV ecosystem. The proposed architecture integrates a hybrid quantum-classical model with SHAP- and LIME-based explainability to enable interpretable trust inference during message authentication and spoofing detection. By leveraging quantum feature embedding through variational circuits, EQAI enhances nonlinear feature discrimination and robustness while maintaining a lightweight design with approximately 4900 trainable parameters and ~19.5 k FLOPs, ensuring real-time inference feasibility on edge devices. Comprehensive experimental validation using the GPS spoofing and CICIoV2024 datasets demonstrates that EQAI achieves a detection accuracy of 92.85%, a Class 3 $F1$-score of 82.1%, and a low false alarm rate (FAR $\leq$ 0.026), outperforming existing ML, BC-based, and FL counterparts. Moreover, the reduced-qubit variant (three-qubit EQAI) shows improved efficiency and stability compared to its two-qubit configuration, confirming the scalability of quantum entanglement in security-critical IoV applications. The explainability analysis using LIME and SHAP highlights the interpretive strength of EQAI in revealing feature-level influences on decision outcomes, thereby promoting transparency and trustworthiness.

The system maintains moderate computational overhead, supports onboard execution, and minimizes communication burden, meeting practical requirements for real-time V2X operations. Future work will focus on deploying EQAI on real CAV platforms using embedded quantum simulators and lightweight processors (e.g., Raspberry Pi and LoRa radios). Additional enhancements will include integrating multimodal sensor fusion (e.g., LiDAR + GNSS), adaptive MAC scheduling, and context-aware message validation via explainable reinforcement learning. Field evaluations are also planned in Gumi, South Korea, to assess performance in real-world traffic. EQAI establishes a foundational step toward deploying quantum-enhanced, explainable, and trustworthy AI for autonomous and connected vehicles. Future work will explore the integration of PQC and BC-assisted model orchestration to build a fully decentralized, resilient, and privacy-preserving quantum trust architecture for next-generation intelligent transportation systems.

References

[1] S. O. Ajakwe and D.-S. Kim, "Facets of security and safety problems and paradigms for smart aerial mobility and intelligent logistics," *IET Intell. Transp. Syst.*, vol. 18, no. S1, pp. 2827–2855, Dec. 2024.

[2] M. Aledhari et al., "Safeguarding connected autonomous vehicle communication: Protocols, intra{-} and inter-vehicular attacks and defenses," *Comput. Secur.*, vol. 151, Apr. 2025, Art. no. 104352.

[3] B. D. Manh, C.-H. Nguyen, D. T. Hoang, and D. N. Nguyen, "Homomorphic encryption-enabled federated learning for privacy-preserving intrusion detection in resource-constrained IoV networks," in *Proc. IEEE 100th Veh. Technol. Conf. (VTC-Fall)*, Oct. 2024, pp. 1–6.

[4] S. O. Ajakwe, K. L. Olabisi, and D.-S. Kim, "Multihop intruder node detection scheme (MINDS) for secured drones' FANET communication," *IET Intell. Transp. Syst.*, vol. 19, no. 1, p. 70080, Jan. 2025.

[5] M. M. Khan, M. Kamal, M. Shabbir, and S. Alahmari, "Enhancing autonomous vehicle security: Federated learning for detecting GPS spoofing attack," *Trans. Emerg. Telecommun. Technol.*, vol. 36, no. 4, p. 70138, Apr. 2025.

[6] V. U. Ihekoronye, S. O. Ajakwe, J. M. Lee, and D.-S. Kim, "DroneGuard: An explainable and efficient machine learning framework for intrusion detection in drone networks," *IEEE Internet Things J.*, vol. 12, no. 7, pp. 7708–7722, Apr. 2025.

[7] S. Sathiyanathn, K. Ranjithkumar, G. Rajesh, V. Revanth, and V. Vignesh, "Blockchain integrated framework to detect and prevent CAV location spoofing attack using GPS time series data learning and quantum cryptography," in *Proc. 2nd Int. Conf. Artif. Intell. Mach. Learn. Appl. Theme, Healthcare Internet Things (AIMLA)*, Mar. 2024, pp. 1–7.

[8] A. M. Aslam, A. Bhardwaj, and R. Chaudhary, "Quantum-resilient blockchain-enabled secure communication framework for connected autonomous vehicles using post-quantum cryptography," *Veh. Commun.*, vol. 52, Apr. 2025, Art. no. 100880.

[9] A. Castiglione and T. Elia, "Securing in-vehicle communications through post-quantum cryptography and network segmentation," *Comput. Electr. Eng.*, vol. 126, Aug. 2025, Art. no. 110488.

[10] N. R. Reddy et al., "Quantumshield-BC: Quantum-secured blockchain framework for enhancing vehicular communications with PQC, QKD and Q-BFT," *Sci. Rep.*, vol. 15, no. 1, 2025, Art. no. 31048, doi: 10.1038/s41598-025-16315-8.

[11] A. Mohd Aslam, R. Chaudhary, and A. Bhardwaj, "An AIoT-enabled digital twin CAVs with a DRL-based framework for trajectory planning," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 10, pp. 18101–18115, Oct. 2025.

[12] P. Babaei, N. Riahinia, O. M. Ebadati E, and A. Azimi, "Towards a data-driven digital twin AI-based architecture for self-driving vehicles," *IET Intell. Transp. Syst.*, vol. 19, no. 1, Jan. 2025, Art. no. e70017.

[13] Z. Li et al., "A survey of reinforcement learning-based motion planning for autonomous driving: Lessons learned from a driving task perspective," 2025, *arXiv:2503.23650*.

[14] U. Ahmad, M. Han, and S. Mahmood, "Enhancing security in connected and autonomous vehicles: A pairing approach and machine learning integration," *Appl. Sci.*, vol. 14, no. 13, p. 5648, Jun. 2024.
[15] S. O. Ajakwe, I. S. Igboanusi, J.-M. Lee, and D.-S. Kim, "IBANDA: A blockchain-assisted defense system for authentication in drone-based logistics," *Drones*, vol. 9, no. 8, p. 590, Aug. 2025.
[16] D. Ranga, A. Rana, S. Prajapat, P. Kumar, K. Kumar, and A. V. Vasilakos, "Quantum machine learning: Exploring the role of data encoding techniques, challenges, and future directions," *Mathematics*, vol. 12, no. 21, p. 3318, Oct. 2024.
[17] S. Choi et al., "Photonic probabilistic machine learning using quantum vacuum noise," *Nature Commun.*, vol. 15, no. 1, p. 7760,Sep. 2024.
[18] M. Fernandez, J. Llorca, and L. Fernandez, "A survey on trust management for VANETs," *ACM Comput. Surveys*, vol. 51, no. 5, pp. 105–112, 2018.
[19] J. Song, T. Liu, X. Chen, and Z. Wu, "Satellite navigation message authentication in GNSS: Research on message scheduler for SBAS L1," *Sensors*, vol. 24, no. 2, 2024, Art. no. 360, doi: 10.3390/s24020360.
[20] T. Leinmuller, E. Schoch, and F. Kargl, "Position verification approaches for vehicular ad hoc networks," *IEEE Wireless Commun.*, vol. 13, no. 5, pp. 16–21, 2006.
[21] Q. Xu, L. Zhang, D. Ou, and W. Yu, "Secure intrusion detection by differentially private federated learning for inter-vehicle networks," *Transportmetrica A, Transp. Sci.*, 2023.
[22] F. Ayaz, Z. Sheng, D. Tian, and Y. L. Guan, "A blockchain based federated learning for message dissemination in vehicular networks," 2021, *arXiv:2109.06667v1*.
[23] W. Liu and P. Papadimitratos, "Self-supervised federated GNSS spoofing detection with opportunistic data," 2025, *arXiv:2505.06171*.
[24] A. Nsour, "Quantum-resilient secure onboard communication for autonomous vehicles using post-quantum cryptography," *Int. J. Comput. Netw. Appl.*, 2025.
[25] M. D. R. Kabir, B. B. Y. Ravi, and S. Ray, "Digital twin technologies for vehicular prototyping: A survey," *IEEE Open J. Intell. Transp. Syst.*, 2025.
[26] G. Aissou, S. Benouadah, H. E. Alami, and N. Kaabouch, "A dataset for GPS spoofing detection on autonomous vehicles," *IEEE DataPort*, 2022, doi: 10.21227/8x3h-2817.
[27] M. D. Firmansyah, I. Rizqa, and F. A. Rafrastara, "Balancing CICIoV2024 dataset with RUS for improved IoV attack detection," *J. Appl. Informat. Comput.*, vol. 9, no. 2, pp. 250–257,Mar. 2025.
[28] S. O. Ajakwe and D. S. Kim, "Explainable quantum-empowered anti-spoofing intelligence for trustworthy connected autonomous vehicles communication," in *Proc. Korena Inst. Commun. Inf. Sci. (KICS) Conf.*, Jun. 2025, pp. 1–2.

Simeon Okechukwu Ajakwe (Senior Member, IEEE) received the Diploma degree in computer science from the Institute of Management Technology, Enugu, Nigeria, in 2004, the Bachelor of Science degree in computer science from Imo State University, Owerri, Nigeria, in 2009, the Master of Science (M.Sc.) degree (Hons.) in information technology from the Federal University of Technology, Owerri, Imo State, Nigeria, in 2016, and the Ph.D. degree (Hons.) in IT-convergence engineering from the Kumoh National Institute of Technology, Gumi, South Korea, in 2023.

He is a Chartered Information Technology Professional (CITP) with many years of IT-related industrial and academic experience. He is a Post-Doctoral Research Fellow at the ICT Convergence Research Center (ICT-CRC), Kumoh National Institute of Technology. He has published several articles and conference papers. His research interests include AI for aerial mobility, trustworthy AI, vehicular networks, the Internet of Medical Things, and information system security.

Dr. Ajakwe is the Association of Computing Machinery (ACM), the Computer Professionals of Nigeria (CPN), and Nigerian Internet Society (NIS). He has received several awards, such as the Best Ph.D. Thesis Award, the Best Paper Award, and the Research Excellence Award. He served as an editor, a guest editor, and a reviewer for reputable peer-reviewed journals.

Dong-Seong Kim (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2003.

He was a full-time Researcher with ERCACI, Seoul National University, from 1994 to 2003. From March 2003 to February 2005, he was a Post-Doctoral Researcher with the Wireless Network Laboratory, School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. From 2007 to 2009, he was a Visiting Professor with the Department of Computer Science, University of California, Davis, Davis, CA, USA. He is currently the Director of the KIT Convergence Research Institute and the ICT Convergence Research Center (ITRC and NRF Advanced Research Center Program) supported by Korean Government at Kumoh National Institute of Technology, Gumi, South Korea. His current research interests are real-time IoT and smart platforms, industrial wireless control networks, networked embedded systems, and fieldbus.

Prof. Kim is a Senior Member of ACM.