

Article

Q-MobiGraphNet: Quantum-Inspired Multimodal IoT and UAV Data Fusion for Coastal Vulnerability and Solar Farm Resilience

Mohammad Aldossary 

Department of Software Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia; mm.aldossary@psau.edu.sa

Abstract

Coastal regions are among the areas most affected by climate change, facing rising sea levels, frequent flooding, and accelerated erosion that place renewable energy infrastructures under serious threat. Solar farms, which are often built along shorelines to maximize sunlight, are particularly vulnerable to salt-induced corrosion, storm surges, and wind damage. These challenges call for monitoring solutions that are not only accurate but also scalable and privacy-preserving. To address this need, Q-MobiGraphNet, a quantum-inspired multimodal classification framework, is proposed for federated coastal vulnerability analysis and solar infrastructure assessment. The framework integrates IoT sensor telemetry, UAV imagery, and geospatial metadata through a Multimodal Feature Harmonization Suite (MFHS), which reduces heterogeneity and ensures consistency across diverse data sources. A quantum sinusoidal encoding layer enriches feature representations, while lightweight MobileNet-based convolution and graph convolutional reasoning capture both local patterns and structural dependencies. For interpretability, the Q-SHAPE module extends Shapley value analysis with quantum-weighted sampling, and a Hybrid Jellyfish–Sailfish Optimization (HJFSO) strategy enables efficient hyperparameter tuning in federated environments. Extensive experiments on datasets from Norwegian coastal solar farms show that Q-MobiGraphNet achieves 98.6% accuracy, and 97.2% F1-score, and 90.8% Prediction Agreement Consistency (PAC), outperforming state-of-the-art multimodal fusion models. With only 16.2 M parameters and an inference time of 46 ms, the framework is lightweight enough for real-time deployment. By combining accuracy, interpretability, and fairness across distributed clients, Q-MobiGraphNet offers actionable insights to enhance the resilience of coastal renewable energy systems.

Keywords: federated learning; multimodal data fusion; quantum-inspired models; coastal vulnerability assessment; solar infrastructure monitoring

MSC: 68T07; 68R10; 68T27; 68W50



Academic Editor: Florin Leon

Received: 20 August 2025

Revised: 9 September 2025

Accepted: 17 September 2025

Published: 22 September 2025

Citation: Aldossary, M.

Q-MobiGraphNet: Quantum-Inspired Multimodal IoT and UAV Data Fusion for Coastal Vulnerability and Solar Farm Resilience. *Mathematics* **2025**, *13*, 3051. <https://doi.org/10.3390/math13183051>

Copyright: © 2025 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Climate change has intensified the occurrence and severity of environmental hazards, with coastal regions facing some of the most significant impacts. Rising sea levels, storm surges, and accelerated erosion threaten not only fragile ecosystems but also critical infrastructure [1]. Among these vulnerable assets, solar farms have become key components in the global shift toward renewable energy, providing clean and sustainable electricity generation [2,3]. However, their frequent placement near coastlines to maximize solar exposure exposes them to harsh environmental stressors such as salt-induced corrosion, high wind

loads, and flooding from extreme weather events [4]. These conditions accelerate wear and degradation, reduce energy efficiency, and create the need for continuous, context-aware monitoring to ensure operational reliability and long-term resilience [5].

Advances in sensing and inspection technologies have greatly improved the ability to monitor such infrastructures. Internet of Things (IoT) sensor networks now offer continuous streams of environmental and operational data—capturing parameters such as temperature, humidity, wind speed, and energy output—providing valuable, real-time insights into system performance [6]. In parallel, uncrewed aerial vehicles (UAVs) have emerged as a flexible and effective tool for high-resolution inspections, capable of detecting structural faults like cracks, hotspots, and corrosion with speed and precision [7]. Despite these individual strengths, IoT- and UAV-based systems are often deployed in isolation, resulting in fragmented datasets that do not fully reflect the interdependencies between environmental stressors and the physical health of infrastructure [8].

The concept of multimodal data fusion—combining heterogeneous data from multiple sources—has gained traction across various fields such as smart city management, disaster response, and environmental hazard detection [9]. Integrating diverse modalities enables richer representations, improves predictive accuracy, and supports more informed decision-making. However, applying multimodal fusion in coastal solar farm monitoring introduces specific challenges. Environmental sensor data and UAV imagery differ in spatial resolution, temporal sampling rates, and statistical properties, creating significant heterogeneity. Coastal environments also change rapidly, demanding fusion strategies that adapt in real time while remaining robust to variability. Furthermore, the distributed nature of solar farms often means that raw data cannot be centralized due to privacy requirements, limited communication bandwidth, or local policy restrictions [10].

Recent advances in deep learning and transformer-based architectures have shown strong potential for handling both structured sensor data and unstructured imagery, with attention mechanisms excelling at modeling cross-modal dependencies [11]. Yet, most existing methods remain limited to single-modality inputs or depend on centralized training, restricting their scalability and flexibility in geographically dispersed monitoring scenarios [12,13]. In addition, many lack explainability, making it difficult for stakeholders to validate outputs and trust automated recommendations—an essential requirement in safety-critical and high-value infrastructure domains.

To address these gaps, this work presents a privacy-preserving multimodal data fusion framework that combines IoT-derived environmental telemetry, UAV-based high-resolution imagery, and geospatial metadata into a unified platform for coastal vulnerability and solar farm infrastructure assessment. The proposed framework employs hybrid feature optimization to align the differing characteristics of each data modality, followed by an attention-guided fusion mechanism that models intricate interdependencies between environmental and visual indicators. The integration of these components within a federated learning environment allows model training to occur across multiple distributed sites without sharing raw data, thereby safeguarding sensitive information. This approach enables accurate, scalable, and interpretable monitoring, offering actionable insights that link environmental hazards to infrastructure health and supporting resilient, sustainable management of coastal energy systems. This study tackles the limitations of fragmented multimodal analysis and the absence of privacy-aware, interpretable frameworks for assessing coastal solar infrastructures. By integrating IoT telemetry, UAV imagery, and geospatial information in a federated environment, the proposed approach reduces data heterogeneity, enhances model transparency, and ensures scalable deployment. The main contributions of this work are summarized as follows:

Research Gaps and Contributions

Despite progress in multimodal learning and coastal infrastructure monitoring, several important research gaps remain:

Research Gaps:

1. **Fragmented analysis:** Most existing studies process IoT telemetry, UAV imagery, and geospatial data separately, which leads to fragmented insights and prevents a holistic understanding of coastal vulnerability.
2. **Heterogeneity and variability:** Differences in sampling rates, spatial resolution, and statistical properties across modalities introduce complexity and make it difficult to achieve seamless multimodal fusion.
3. **Scalability and privacy limitations:** Many existing approaches rely on centralized training, which does not scale well and is unsuitable for distributed monitoring scenarios where privacy and bandwidth constraints are critical.
4. **Limited interpretability:** A large number of multimodal models still function as black boxes, offering limited transparency and making it hard for stakeholders to validate predictions in high-stakes applications.
5. **Inefficient hyperparameter tuning:** Conventional tuning methods are often slow, unstable, and computationally expensive in federated environments, which hampers convergence and practical deployment.

Contributions:

1. **Introduction of Q-MobiGraphNet**, a quantum-inspired multimodal framework that unifies IoT data, UAV imagery, and geospatial information to overcome fragmentation and capture both spatial and temporal dependencies.
2. **Development of the Multimodal Feature Harmonization Suite (MFHS)**, a preprocessing pipeline that standardizes, synchronizes, and aligns diverse modalities, reducing heterogeneity and enabling reliable multimodal fusion in federated learning settings.
3. **Design of the Q-SHAPE explainability module**, which extends Shapley value analysis with quantum-weighted sampling to provide transparent, interpretable, and stable explanations of feature importance.
4. **Proposal of the Hybrid Jellyfish–Sailfish Optimization (HJFSO) algorithm**, which balances exploration and exploitation to achieve faster convergence, improved stability, and lower computational cost compared with traditional methods.
5. **Deployment of a privacy-preserving monitoring framework** for coastal renewable infrastructures, ensuring scalability, interpretability, and resource efficiency in managing solar farms under climate-induced environmental stressors.

The remainder of this article is structured as follows. Section 2 reviews related work on multimodal data fusion, coastal monitoring, and infrastructure assessment, with an emphasis on current advances and unresolved challenges. Section 3 describes the proposed methodology, including the harmonization pipeline, the Q-MobiGraphNet architecture, and the optimization strategies for federated deployment. Section 4 presents experimental results, comparative evaluations, ablation studies, and interpretability analyses that demonstrate the effectiveness of the framework. Section 5 concludes the paper and discusses future research directions aimed at enhancing scalability, robustness, and practical applicability in coastal renewable infrastructure monitoring.

2. Related Work

Multimodal data fusion improves environmental threat assessment and infrastructure monitoring. IoT-based environmental sensors, UAV imaging, and geospatial metadata provide a fuller situational picture and enable informed decision-making in coastal risk

and solar farm infrastructure contexts. Many single-modality systems fail to reflect the complex relationships between environmental dynamics, structural integrity, and spatial risks, especially in quickly changing coastal zones. Recent research has used deep learning, transformer-based fusion, and cross-domain feature alignment to merge time-series sensor data with high-resolution UAV imagery to address data heterogeneity, spatiotemporal variability, and distributed deployment constraints.

DeepLCZChange [14] uses a deep change-detection network to examine urban cooling impacts using LiDAR and Landsat surface temperature data. The method enhances climate resilience mapping by incorporating geographic specificity through the collection of vegetation-driven microclimate trends. While successful for urban analysis, it lacks IoT telemetry, UAV photography, and coastal hazard data, limiting its applicability for monitoring infrastructure risks. However, it highlights the benefits of multimodal geospatial fusion and the necessity of IoT time-series inputs and UAV-based coastline and solar farm inspections. The taxonomy in [15] categorizes deep fusion strategies into feature-, alignment-, contrast-, and generation-based categories, focusing on urban data integration, a helpful framework for choosing optimal modality fusion approaches. The study does not cover infrastructure monitoring or environmental threat forecasts. A Hybrid Feature Selector and Extractor (HyFSE) is used in the proposed methodology to efficiently integrate IoT time series, UAV imagery, and geospatial metadata for federated learning-based coastal and solar infrastructure assessment.

In [16], an innovative flood resilience platform is created to alert to potential hazards using community-scale and infrastructure-mounted sensors. The system responds swiftly to localized flooding but lacks UAV-based damage assessments and solar asset monitoring. A federated framework for multi-site privacy and scalability is missing. UAV imaging and IoT-driven solar performance indicators are combined into a privacy-conscious multimodal model for hazard detection and infrastructure health evaluation in this work. The author in [17] uses a Transformer-based architecture to combine multimodal time-series and image data for smart city applications. Attention layer models inter-modal interactions, improving prediction accuracy. UAV–IoT synchronization, cross-modal feature selection for coastal or solar applications, and federated learning compatibility are not supported. HyFSE-based feature refinement and federated deployment enable scalable operations across geographically scattered infrastructure sites in the proposed system.

The LiDAR, radar, and optical data in the LRVFNet model [18] enable occlusion-robust object detection in autonomous driving. Attention processes boost performance under challenging situations. Strong multimodal integration, but with a focus on vehicle perception rather than infrastructure health. Attention-guided fusion to integrate IoT data, UAV imagery, and geographic features is used to predict coastal vulnerability and evaluate solar farm performance. In [19], MMFnet blends altimeter and scatterometer data with ConvLSTM and LSTM algorithms to predict sea level anomalies. Integration of several sources improves temporal forecasting. Although important to oceanography, it lacks UAV-based structural inspection and IoT-derived infrastructure operational data. This is expanded by integrating environmental forecasts into a coastal infrastructure and solar farm risk assessment framework in a federated learning architecture.

UAV RGB images and a CNN-SVM pipeline are used in [20] to classify multi-class PV defects. The method's accuracy across various defect types demonstrates UAV imagery's potential for asset monitoring. Its visual-only approach restricts diagnostic insight. The proposed design in this work combines UAV-based fault detection with IoT performance indicators to better understand solar farm health under environmental stressors. In [21], an AE-LSTM method detects anomalies in PV telemetry without labeled datasets, making it ideal for early warnings. UAV visual verification and geospatial context are needed to

locate and diagnose issues. The proposed technology in this work couples IoT anomaly signals with UAV photography and spatial mapping to precisely locate coastal solar faults.

In [22], texture-based infrared thermography image analysis, GLCM features, and SVM classification are used to detect PV faults. This interpretable technique identifies faults but lacks deep learning scalability and multimodal integration. Explainable AI and multimodal fusion of IoT, UAV, and GIS data ensure interpretability and infrastructure assessment in this approach. In [23], the authors train a U-Net and a DeepLabV3+ on UAV thermal imagery to detect high-Dice PV faults. Despite accurate segmentation, the approach only uses pictures and ignores environmental or operational data. In this work, the segmentation and IoT-based performance analysis provide integrated assessments for speedy repairs and long-term resilience planning. The CNN model in [24] is ideal for UAV-based PV defect detection, providing excellent classification throughput. The diagnostic scope is limited by single-modality use. Context-aware insights for coastal solar farms are supplied by merging IoT operational data and hazard indicators. In [25], a lightweight CNN with transfer learning achieves near-perfect accuracy in PV defect classification with minimal computational cost. However, multi-site flexibility and federated learning are lacking. These efficiency gains are used in a federated training paradigm to enable privacy-preserving adaptation across multiple locales. In distribution networks, cooperative AC/DC control strategies have been proposed to mitigate voltage violations [26], highlighting the importance of coordinated optimization in energy systems.

The authors in [27] proposed a semi-supervised GAN-based ppFDetector that accurately identifies abnormal PV states without substantial labeling. Its imagery-only approach reduces context. The proposed solution in this work correlates abnormalities with environmental and operational characteristics using UAV and IoT streams. According to [28], the ICNM model is optimized for real-time fault detection, considering its speed and accuracy. UAV-based validation and multimodal capabilities are missing. In this work, the HyFSE pipeline blends the most useful UAV, IoT, and GIS elements for efficient, comprehensive monitoring. Skip-connected SC-DeepCNN in [25] enhances PV imagery hotspot localization, but relies on manual region recommendations and disregards IoT data. Instead of manual assessment, the proposed system uses automated cross-modal fusion to evaluate flaws and operational consequences. The refined VGG-16 model in [29] performs well in supervised PV fault classification but is limited to centralized training and lacks privacy protections. In this work, the federated learning system solves these problems while maintaining the accuracy of coastal and solar infrastructure detection.

Multimodal fusion has advanced in urban analytics, autonomous driving, photovoltaic defect detection, and oceanographic forecasting. As shown in Table 1, most techniques are limited by single-modality, domain-specific restrictions, or the lack of integrated UAV-IoT-geospatial pipelines for coastal and solar infrastructure monitoring. Federated learning paradigms, which protect data privacy across geographically dispersed sites, are rarely used in these works. This gap highlights the need for a unified, privacy-aware framework that can seamlessly integrate disparate data sources and meet the spatiotemporal complexity and operational needs of coastal hazard assessment and solar farm asset appraisal. The proposed solution addresses these deficiencies by leveraging cross-modal feature selection, explainable AI, and scalable federated deployment to improve multimodal fusion techniques.

Table 2 presents a side-by-side comparison of Q-MobiGraphNet with recent state-of-the-art methods. The results clearly illustrate how the proposed framework stands out by offering broader modality integration, improved interpretability, higher computational efficiency, and strong support for privacy-preserving scalability.

Table 1. Literature on multimodal IoT and UAV data fusion for coastal and solar infrastructure assessment.

Ref.	Objective	Method	Key Achievements	Identified Limitations
[14]	Improve urban cooling analysis using multiple geospatial data sources	LiDAR and Landsat surface temperature fusion via deep change detection network	Delivers fine-grained spatial maps and reveals vegetation impacts on microclimate resilience	Does not integrate IoT sensors, UAV imagery, or coastal hazard indicators
[15]	Establish a structured taxonomy for deep multimodal fusion	Categorization into feature-, alignment-, contrast-, and generation-based fusion types	Provides a clear reference framework for modality integration strategies	Focused on urban datasets only; no infrastructure or federated learning context
[16]	Build a community-scale flood resilience prediction system	IoT and infrastructure sensors for real-time hazard alerts	Offers timely and localized flood warnings with high responsiveness	Lacks UAV-based visual validation and solar infrastructure monitoring
[17]	Advance urban analytics with multimodal learning	Transformer-based fusion of time series and imagery using attention mechanisms	Captures cross-modal dependencies, improving prediction performance	No UAV–IoT synchronization or adaptation to coastal/solar domains
[18]	Boost detection reliability in autonomous driving	LiDAR–radar–vision fusion with attention enhancement	Maintains accuracy in complex and adverse driving conditions	Tailored for automotive use; not suitable for infrastructure risk assessment
[19]	Predict sea level anomalies from oceanographic data	ConvLSTM–LSTM fusion of altimeter and scatterometer measurements	Enhances forecasting accuracy for marine environments	Excludes UAV imagery and IoT-based monitoring metrics
[20]	Classify PV defects using aerial imagery	UAV RGB imagery with CNN–SVM fusion	Achieves high defect classification accuracy across PV categories	Limited to visual data without environmental integration
[21]	Detect PV performance anomalies from telemetry	AE–LSTM applied to time-series sensor data	Enables unsupervised detection without labeled datasets	No UAV visual validation or hazard mapping
[22]	Identify PV faults via thermal imagery	GLCM texture feature extraction with SVM	Produces interpretable fault detection from thermal patterns	Does not incorporate multimodal data sources
[23]	Localize PV defects in UAV thermal images	U-Net and DeepLabV3+ segmentation architectures	Attains high segmentation accuracy with strong Dice scores	Restricted to imagery without IoT/environmental context
[24]	Large-scale PV defect detection from UAV images	CNN optimized for efficient classification	Supports scalable, high-throughput inspection workflows	Relies on a single modality, omitting IoT and hazard indicators
[25]	Efficient PV defect classification	Lightweight CNN with transfer learning	Balances strong accuracy with low computation needs	Lacks multi-site adaptation or federated training
[27]	Semi-supervised PV fault detection	GAN-based reconstruction of normal operating states	Effective in low-label scenarios	Limited to image-only analysis
[28]	Real-time PV fault identification	ICNM model for speed–accuracy optimization	Suitable for real-time detection with solid accuracy	No multimodal fusion or UAV confirmation
[25]	Enhance PV hotspot localization	SC-DeepCNN with skip connections	Improves hotspot detection in PV modules	Dependent on handcrafted ROI inputs; no IoT integration
[29]	Supervised PV fault classification	Fine-tuned VGG-16 on labeled PV datasets	Strong classification performance	Centralized setup without privacy-preserving mechanisms

Table 2. Comparative analysis of Q-MobiGraphNet with recent state-of-the-art multimodal fusion methods.

Method	Modalities Used	Learning Setup	Interpretability	Optimization Strategy	Accuracy (%)	Params (M)	Inference Time (ms)	Scalability/Privacy
DCDN [14]	LiDAR + Landsat	Centralized	No	Standard SGD	80.5	45.8	97	No
Transformer-based fusion [17]	Time-series + imagery	Centralized	Partial (attention)	Standard SGD	82.1	38.5	83	No
LiDAR–radar–vision fusion [18]	LiDAR + radar + vision	Centralized	Limited	Gradient-based	83.6	34.2	79	No
CNN–SVM fusion [20]	UAV RGB imagery	Centralized	No	Manual tuning	86.2	26.4	69	No
AE–LSTM [21]	PV telemetry only	Centralized	No	Autoencoder init.	87.0	23.8	65	No
U-Net + DeepLabV3+ [23]	UAV thermal imagery	Centralized	No	Deep CNN	89.5	21.7	61	No
Lightweight CNN + TL [25]	UAV RGB imagery	Centralized	No	Transfer learning	91.0	14.9	48	No
GAN-based reconstruction [27]	UAV imagery only	Centralized	Limited	GAN-based training	92.0	22.1	63	No
VGG-16 (fine-tuned) [29]	UAV imagery only	Centralized	No	Gradient descent	88.7	138.3	101	No
Proposed Q-MobiGraphNet	IoT + UAV + GIS	Federated	Yes (Q-SHAPE)	Hybrid Jellyfish–Sailfish (HJFSO)	98.6	16.2	46	Yes (scalable, privacy-preserving)

3. Proposed Methodology

This section presents the proposed framework, Q-MobiGraphNet, which is developed as a quantum-inspired, multimodal, and federated classification model for coastal vulnerability prediction and solar infrastructure assessment. The methodology follows a structured process that begins with multimodal data collection and preprocessing, supported by the Multimodal Feature Harmonization Suite (MFHS) for standardizing and aligning IoT telemetry, UAV imagery, and geospatial metadata. Once harmonized, the inputs are processed through the Q-MobiGraphNet architecture, which combines quantum sinusoidal encoding, lightweight MobileNet-based convolution, and graph convolutional reasoning to capture spatiotemporal and structural dependencies across different modalities effectively. To strengthen transparency, the Q-SHAPE module provides quantum-weighted Shapley-based feature attribution, while convergence stability is achieved through the Hybrid Jellyfish–Sailfish Optimization (HJFSO) strategy for hyperparameter tuning. The pictorial view of the proposed model is shown in Figure 1. The following subsections describe each stage of the methodology in detail, along with their mathematical formulations, highlighting how the framework enables privacy-preserving, scalable, and interpretable monitoring in federated environments.

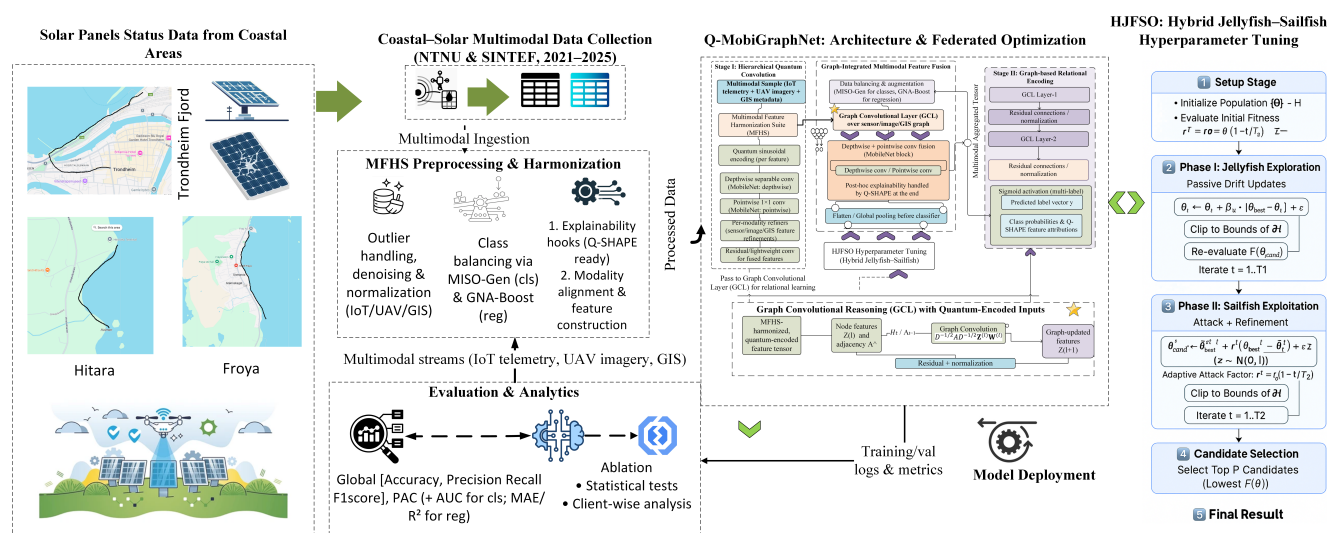


Figure 1. Proposed framework for coastal vulnerability prediction and solar infrastructure assessment.

3.1. Data Collection and Preprocessing

The dataset used in this study was developed through a joint effort between the Norwegian University of Science and Technology (NTNU) and SINTEF Digital. Data were collected over four years, from January 2021 to February 2025, focusing on coastal areas such as Trondheim Fjord, Hitra, and Frøya in central Norway, where pilot-scale coastal solar farms have been deployed [30]. Environmental variables were recorded through IoT-enabled sensor grids installed by SINTEF, while UAV-based aerial surveys conducted by NTNU’s Department of Engineering Cybernetics provided high-resolution imagery. To ensure reliability and consistency across the different data modalities, all records were anonymized and processed through a standardized preprocessing pipeline. The final dataset, curated and managed by the NTNU Smart Energy and Infrastructure Lab, is hosted on a controlled Kaggle repository. This dataset underpins the modeling of coastal vulnerability and the performance of solar infrastructure under diverse environmental conditions. A detailed overview of dataset features is presented in Table 3.

Table 3. Summary of dataset feature categories and variables.

Feature Category	Features
Environmental IoT Sensor Data	Temperature, Humidity, Wind Speed, Solar Irradiance
UAV-Derived Image Attributes	Panel Damage Score, Shadow Coverage, Glare Ratio
Solar Infrastructure Metrics	Panel Voltage, Current, Power Output, Operational Status
GIS and Topographic Data	Elevation, Slope, Distance from Coast, Land Cover Type
Metadata and Fusion Support	Data Source, Sensor ID, Modality Type, Region Type
Derived Analytical Features	Panel Efficiency Ratio, Risk Zone Label, Anomaly Score
Target Labels	Coastal Vulnerability Level, Flood Risk Score, Panel Health Status, Energy Output Efficiency

3.2. Preprocessing and Feature Harmonization Pipeline

To effectively integrate heterogeneous inputs from IoT sensors, UAV imagery, and contextual metadata into Q-MobiGraphNet, a nine-stage preprocessing and harmonization pipeline is designed. This pipeline ensures consistency across modalities, enhances robustness against noise, and prepares the data for federated learning [31,32]. The steps are shown in Algorithm 1.

Algorithm 1 Preprocessing and feature harmonization pipeline.

```

1: Input: Raw multimodal dataset  $\mathcal{D} = \{\mathbf{S}, \mathcal{I}, \mathbf{M}, \mathbf{C}, \mathcal{T}_s, \mathcal{T}_u\}$ 
2:    $\mathbf{S} \in \mathbb{R}^{N \times F}$ : IoT sensor matrix;  $\mathcal{I}$ : UAV images;  $\mathbf{M}$ : geospatial features;  $\mathbf{C}$ : metadata
3:    $\mathcal{T}_s$ : sensor timestamps;  $\mathcal{T}_u$ : UAV frame timestamps
4: Output: Harmonized matrix  $\mathbf{F} \in \mathbb{R}^{N \times D}$  and per-sample vectors  $\{\mathbf{f}_n\}_{n=1}^N$ 
5: Step 1: Z-score normalization (Equation (1))
6: for  $f = 1$  to  $F$  do
7:    $\mu_f \leftarrow \text{mean}(s_{:,f}), \sigma_f \leftarrow \text{std}(s_{:,f})$ 
8:    $s_{n,f} \leftarrow (s_{n,f} - \mu_f) / \sigma_f \forall n$ 
9: end for
10: Step 2: Range scaling (Equation (2))
11: for each bounded feature  $f$  do
12:    $a_f \leftarrow \min(s_{:,f}), b_f \leftarrow \max(s_{:,f})$ 
13:    $s_{n,f} \leftarrow (s_{n,f} - a_f) / (b_f - a_f) \forall n$ 
14: end for
15: Step 3: UAV preprocessing (Equation (3))
16: for each  $I_n \in \mathcal{I}$  do
17:   Resize to  $256 \times 256$ , grayscale, equalize
18:   Extract  $\mathbf{v}_n \leftarrow [\phi_1, \phi_2, \phi_3, \phi_4]$ 
19: end for
20: Step 4: One-hot encode metadata (Equation (4))
21:  $\mathbf{c}_n \leftarrow \text{OneHot}(\mathbf{C}_n)$ 
22: Step 5: Temporal sync (Equation (5))
23:  $t_n^{\text{sync}} \leftarrow \arg \min_{t_u^{(m)}} |t_s^{(n)} - t_u^{(m)}|$ 
24: Step 6: Missing values (Equations (6) and (7))
25: Interpolate continuous; mode-impute categorical
26: Step 7: Low-variance removal (Equation (8))
27: Discard  $f$  if  $\text{Var}(s_{:,f}) < \epsilon$ 
28: Step 8: Balancing (Equations (9) and (10))
29: Apply MISO-Gen (classification) or GNA-Boost (regression)
30: Step 9: Fusion (Equation (11))
31:  $\mathbf{f}_n \leftarrow [\mathbf{s}_n | \mathbf{v}_n | \mathbf{m}_n | \mathbf{c}_n]$ 
32:  $\mathbf{F} \leftarrow [\mathbf{f}_1; \dots; \mathbf{f}_N]$ 
33: Return  $\mathbf{F}$  and  $\{\mathbf{f}_n\}_{n=1}^N$ 

```

Initially, by normalizing environmental sensor readings $\mathbf{S} \in \mathbb{R}^{N \times F}$, where N is the number of samples and F is the number of features [33]. Z-score normalization was applied to reduce scale bias and stabilize optimization:

$$s_{n,f}^{\text{norm}} = \frac{s_{n,f} - \mu_f}{\sigma_f}, \quad (1)$$

where μ_f and σ_f are the mean and standard deviation of feature f . This ensured that all features contributed equally during training. Next, features with natural operational bounds (e.g., irradiance, voltage, current) were rescaled to the $[0, 1]$ range using gradient-aware range indexing:

$$s_{n,f}^{\text{scale}} = \frac{s_{n,f} - \min(s_{:,f})}{\max(s_{:,f}) - \min(s_{:,f})}. \quad (2)$$

This preserved relative magnitudes over time while allowing direct comparison across modalities. For UAV imagery, each RGB frame $I_n(u, v, c)$ was resized to 256×256 , converted to grayscale, and enhanced using histogram equalization [34,35]:

$$I_n^{\text{eq}}(u, v) = \left\lfloor \frac{(L-1)}{H \cdot W} \sum_{k=0}^{I_n(u,v)} h(k) \right\rfloor, \quad (3)$$

where $h(k)$ denotes histogram counts, and L , H , and W represent grayscale levels and image dimensions. From these enhanced images, descriptors such as vegetation coverage (ϕ_1), degradation score (ϕ_2), tilt index (ϕ_3), and shadow occlusion (ϕ_4) were extracted and embedded into a UAV-specific vector $\mathbf{v}_n$. Categorical metadata (e.g., terrain class, panel type) were transformed into one-hot encodings:

$$\chi_{n,k} = \begin{cases} 1, & \text{if sample } n \text{ belongs to class } k, \\ 0, & \text{otherwise,} \end{cases} \quad k \in \{1, \dots, \kappa\}. \quad (4)$$

This retained semantic meaning while making the data machine-readable. To synchronize modalities with different sampling rates, a nearest-frame alignment scheme was applied. For each sensor timestamp $t_s^{(n)}$, the closest UAV frame $t_u^{(m)}$ was selected:

$$t_n^{\text{sync}} = \arg \min_{t_u^{(m)}} |t_s^{(n)} - t_u^{(m)}|, \quad (5)$$

Ensuring temporal coherence between sensor and image data. Missing values were then addressed. Continuous gaps were filled with linear interpolation [36]:

$$s_{n,f}^{\text{fill}} = \frac{s_{n-1,f} + s_{n+1,f}}{2}, \quad \text{if } s_{n,f} \text{ is missing,} \quad (6)$$

While categorical gaps were imputed using mode-based inference:

$$\chi_n^{\text{fill}} = \arg \max_{c \in \mathcal{C}} \text{count}(\chi_{:,f} = c), \quad (7)$$

where $\mathcal{C}$ is the set of possible categories. Low-variance features were removed to reduce redundancy and improve interpretability. Any feature with variance below the threshold ϵ was discarded:

$$\text{Var}(s_{:,f}) < \epsilon \Rightarrow \text{discard feature } f. \quad (8)$$

This step improved sparsity and reduced noise in the feature space. To address class imbalance, two complementary strategies were used [37,38]. For classification, the Minority Synthetic Oversampling Generator (MISO-Gen) created synthetic minority samples:

$$\mathbf{s}_{\text{new}} = \mathbf{s}_i + \lambda \cdot (\mathbf{s}_{\text{NN}} - \mathbf{s}_i), \quad \lambda \sim \mathcal{U}(0, 1), \quad (9)$$

where $\mathbf{s}_i$ is a minority sample and $\mathbf{s}_{\text{NN}}$ its nearest neighbor. For regression tasks, Gaussian Noise Augmentation (GNA-Boost) introduced controlled perturbations:

$$\hat{y}_n = y_n + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2). \quad (10)$$

This balanced the dataset while improving generalization. Finally, all modality-specific features were concatenated into a single multimodal representation:

$$\mathbf{f}_n = [\mathbf{s}_n \mid \mathbf{v}_n \mid \mathbf{m}_n \mid \mathbf{c}_n], \quad (11)$$

where $\mathbf{s}_n$ denotes normalized sensor features, $\mathbf{v}_n$ UAV-derived metrics, $\mathbf{m}_n$ geospatial metadata, and $\mathbf{c}_n$ contextual encodings. The resulting matrix $\mathbf{F} \in \mathbb{R}^{N \times D}$ provided the harmonized multimodal input for Q-MobiGraphNet.

3.3. Proposed Classification Model: Q-MobiGraphNet

To enable robust and interpretable anomaly detection in underwater drone surveillance, Q-MobiGraphNet is introduced. This quantum-inspired hybrid classification architecture integrates temporal, spatial, and contextual learning across multimodal data sources, as shown in Figure 2. The model begins with a quantum sinusoidal encoding layer, which transforms each feature f_k from the preprocessed input vector $\mathbf{f}_n \in \mathbb{R}^D$ into a phase-based signal using a dual sinusoidal function. This representation, designed to capture periodicity and nonlinearity, is defined as:

$$Q_n^{(k)} = \sin\left(\frac{f_k}{\alpha^{2k/D}}\right) + \cos\left(\frac{f_k}{\alpha^{2k/D}}\right), \quad (12)$$

where α is a scaling constant (empirically set to 10,000), and D denotes the feature dimensionality. This encoding enhances the expressiveness of both continuous and categorical data by injecting quantum-like diversity into the feature space.

The encoded feature set is then passed through a depthwise separable convolutional module inspired by MobileNet to extract local spatial patterns. The depthwise convolution is applied independently to each input channel, enabling lightweight filtering

$$\mathbf{H}_d^{(i)} = \sigma\left(\mathbf{W}_d^{(i)} * Q_n^{(i)} + b_d^{(i)}\right), \quad (13)$$

where $*$ denotes convolution, $\mathbf{W}_d^{(i)}$ is the kernel for channel i , and $\sigma(\cdot)$ represents the Swish activation function:

$$\sigma(z) = z \cdot \frac{1}{1 + e^{-z}}. \quad (14)$$

This is followed by a pointwise 1×1 convolution to recombine feature maps across channels:

$$\mathbf{H}_p = \phi(\mathbf{W}_p \cdot \mathbf{H}_d + b_p), \quad (15)$$

where ϕ is again the Swish function, and $\mathbf{W}_p$ represents the pointwise weights. Together, these two steps balance computational efficiency with feature richness.

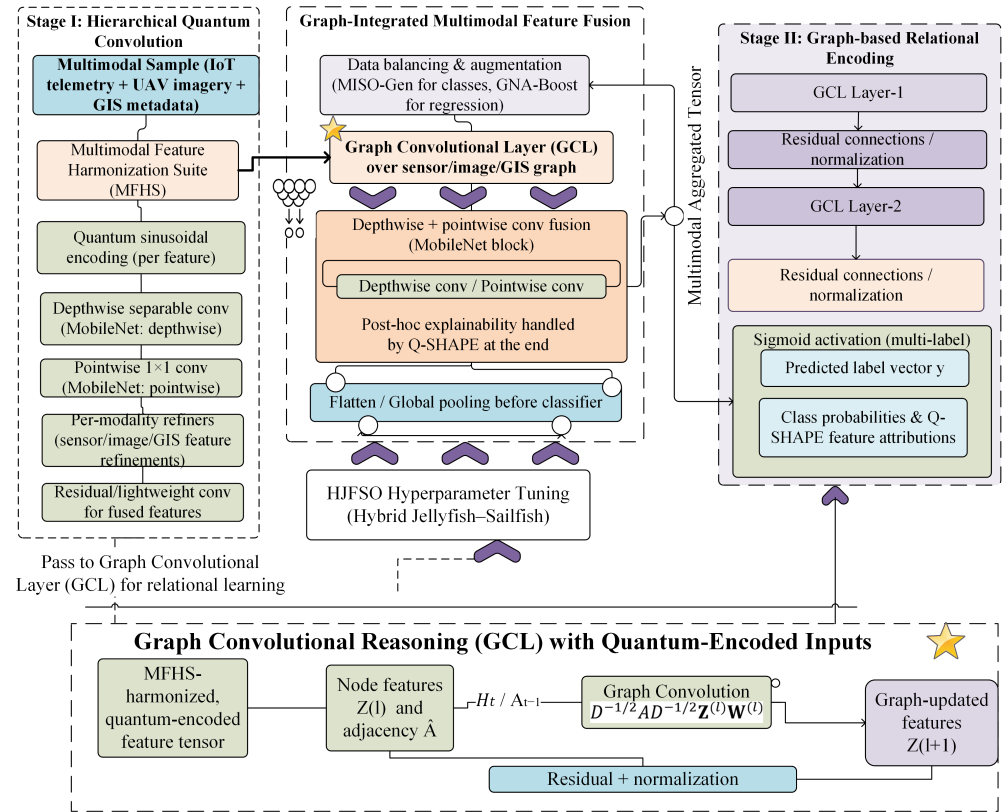


Figure 2. Proposed Q-MobiGraphNet architecture. Star marks module usage/expansion, while circles represent feature maps, $\hat{A}$ is the adjacency matrix with self-loops.

To model structural relationships across spatial or temporal domains—such as UAV image tiles, sensor node interactions, or temporal sequences—a graph convolutional layer (GCL) is employed. Here, the convolution operates over a constructed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each node represents a learned segment, and edges encode neighborhood dependencies. The GCL propagates feature information using:

$$\mathbf{Z}^{(l+1)} = \rho\left(\hat{D}^{-1/2} \hat{A} \hat{D}^{-1/2} \mathbf{Z}^{(l)} \mathbf{W}^{(l)}\right), \quad (16)$$

where $\hat{A} = A + I$ includes self-loops, $\hat{D}$ is the degree matrix of $\hat{A}$, $\mathbf{W}^{(l)}$ denotes learnable weights at layer l , and $\rho(\cdot)$ is the LeakyReLU function. This enables the model to learn higher-order interactions beyond local convolutions.

The graph-encoded features are then flattened into a single vector $\mathbf{z}$ and passed through a fully connected classification layer to compute the raw logits:

$$\mathbf{o} = \mathbf{W}_c \cdot \mathbf{z} + \mathbf{b}_c, \quad (17)$$

where $\mathbf{W}_c$ and $\mathbf{b}_c$ denote the weights and biases of the classifier. For multi-label outputs, a sigmoid activation is applied independently to each class score:

$$\hat{y}_i = \frac{1}{1 + e^{-o_i}}. \quad (18)$$

The model is trained using the binary cross-entropy loss, which penalizes incorrect predictions across all output dimensions:

$$\mathcal{L} = - \sum_{i=1}^C [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (19)$$

where C is the number of target classes and y_i is the ground truth label for class i .

To ensure interpretability of predictions, Q-MobiGraphNet integrates Q-SHAPE, a quantum-inspired explainability module that extends the concept of Shapley values through quantum-weighted sampling. Each feature's contribution ϕ_k is estimated using:

$$\phi_k = \sum_{S \subseteq F \setminus \{k\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f(S \cup \{k\}) - f(S)], \quad (20)$$

where $f(S)$ denotes the model output using only feature subset S . Q-SHAPE approximates this formulation using simulated amplitude encoding and phase-based sampling to prioritize features under quantum principles efficiently. The steps of the proposed method are shown in Algorithm 2.

Algorithm 2 Q-MobiGraphNet: quantum-driven multimodal classification framework.

```

1: Input: Preprocessed feature vector  $\mathbf{f}_n \in \mathbb{R}^D$ 
2: Output: Predicted label vector  $\hat{\mathbf{y}}_n$ 
3: // Quantum Sinusoidal Encoding
4: for each feature  $f_k$  in  $\mathbf{f}_n$  do
5:    $Q_n^{(k)} \leftarrow \sin\left(\frac{f_k}{a^{2k/D}}\right) + \cos\left(\frac{f_k}{a^{2k/D}}\right)$ 
6: end for
7: // Depthwise Separable Convolution (MobileNet)
8: for each channel  $i$  in  $Q_n$  do
9:    $\mathbf{H}_d^{(i)} \leftarrow \text{Swish}(\mathbf{W}_d^{(i)} * Q_n^{(i)} + b_d^{(i)})$ 
10: end for
11:  $\mathbf{H}_p \leftarrow \text{Swish}(\mathbf{W}_p \cdot \mathbf{H}_d + b_p)$ 
12: // Graph Convolution for Relational Learning
13: Construct graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with adjacency matrix  $A$ 
14:  $\hat{A} \leftarrow A + I, \hat{D} \leftarrow \text{diag}(\sum_j \hat{A}_{ij})$ 
15:  $\mathbf{Z}^{(0)} \leftarrow \text{reshape}(\mathbf{H}_p)$ 
16: for each GCN layer  $l$  do
17:    $\mathbf{Z}^{(l+1)} \leftarrow \text{LeakyReLU}\left(\hat{D}^{-1/2} \hat{A} \hat{D}^{-1/2} \mathbf{Z}^{(l)} \mathbf{W}^{(l)}\right)$ 
18: end for
19: // Fully Connected Classification
20:  $\mathbf{z} \leftarrow \text{flatten}(\mathbf{Z}^{(L)})$ 
21:  $\mathbf{o} \leftarrow \mathbf{W}_c \cdot \mathbf{z} + \mathbf{b}_c$ 
22: for each class  $i$  do
23:    $\hat{y}_i \leftarrow \frac{1}{1 + \exp(-o_i)}$ 
24: end for
25: // Binary Cross-Entropy Loss (Training Phase)
26:  $\mathcal{L} \leftarrow -\sum_{i=1}^C [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$ 
27: // Explainability via Q-SHAPE
28: for each feature  $k$  do
29:   Approximate  $\phi_k$  using quantum-weighted Shapley sampling:
30:    $\phi_k \leftarrow \sum_{S \subseteq F \setminus \{k\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f(S \cup \{k\}) - f(S)]$ 
31: end for
32: Return:  $\hat{\mathbf{y}}_n, \mathcal{L}$ , and feature contributions  $\{\phi_k\}$ 

```

In essence, Q-MobiGraphNet introduces a unified, interpretable, and edge-efficient framework by combining quantum-encoded inputs, a lightweight convolutional design, structural graph reasoning, and explainable feature attribution. This makes it highly effective for real-time, privacy-preserving underwater surveillance scenarios requiring high accuracy and trustworthiness.

3.4. Parameter Tuning with HJFSO

Parameter selection is one of the most important factors influencing the performance of metaheuristic algorithms. Common strategies for setting values such as population size,

learning rate, or exploration–exploitation weights usually fall into three categories. The first is manual tuning or grid search, where parameters are adjusted through trial and error. While simple, this method can be time-consuming and computationally expensive. The second is relying on fixed defaults suggested in earlier studies, which may work in some cases but often fail to generalize across different problem domains. The third is adaptive or self-adaptive schemes, where parameters are adjusted automatically as the optimization progresses. Beyond these, more advanced techniques such as hyper-heuristics or meta-optimization use one algorithm to tune another, but this tends to add significant complexity and overhead. To address these limitations, our work uses the Hybrid Jellyfish–Sailfish Optimization (HJFSO). This method adaptively balances exploration and exploitation, updating parameters in response to population diversity rather than static rules or manual choices, resulting in a more stable and efficient optimization process.

The performance of the proposed Q-MobiGraphNet for multimodal coastal vulnerability assessment and solar infrastructure evaluation is influenced not only by the strength of its feature extraction and graph reasoning components but also by the precise selection of its hyperparameters. While the Hybrid Feature Selector and Extractor (HyFSE) module reduces redundancy and ensures a compact, discriminative input space, the overall learning process still relies on key parameters such as the learning rate η , batch size B , dropout probability p_d , number of graph convolutional layers L_g , quantum sinusoidal encoding scaling factor λ_q , attention head count H_a , and weight decay coefficient ω_d . If these parameters are not optimally chosen, the model may experience slow convergence, overfitting, or underutilization of its representational capacity, especially under heterogeneous federated learning conditions.

To navigate this complex, high-dimensional hyperparameter space efficiently, a Hybrid Jellyfish–Sailfish Optimization (HJFSO) strategy is proposed [39,40]. This hybrid method combines the exploratory ability of the Jellyfish Search Optimizer (JSO) with the refinement-focused behavior of the Sailfish Optimizer (SFO). The process alternates between broad exploration of the search space and focused exploitation of promising regions, with the transition between these phases adaptively guided by the diversity of the candidate population.

During the exploration phase, candidate hyperparameter configurations $\theta_i^{(t)}$ simulate jellyfish movement in ocean currents, drifting either passively or actively. The passive drift update rule is expressed as

$$\theta_i^{(t+1)} = \theta_i^{(t)} + \beta \cdot \mathcal{U}(-1, 1) \cdot \left| \theta_i^{(t)} - \theta_{\text{best}}^{(t)} \right|, \quad (21)$$

where β is the drift intensity, $\mathcal{U}(-1, 1)$ is a uniformly distributed random number in the range $[-1, 1]$, and $\theta_{\text{best}}^{(t)}$ is the best configuration discovered so far. This mechanism ensures exploration is both stochastic and biased toward high-performing solutions.

In the exploitation phase, the approach emulates the hunting tactics of sailfish attacking sardine swarms, where candidate solutions move aggressively toward the best-known solution with small perturbations to maintain diversity and avoid premature convergence:

$$\theta_i^{(t+1)} = \theta_{\text{best}}^{(t)} + r \cdot \left(\theta_{\text{best}}^{(t)} - \theta_i^{(t)} \right) + \epsilon \cdot \mathcal{N}(0, 1), \quad (22)$$

where r is the adaptive attack factor, ϵ controls the perturbation magnitude, and $\mathcal{N}(0, 1)$ denotes Gaussian noise for fine-grained search adjustments.

The population diversity metric determines the decision to explore or exploit:

$$\delta^{(t)} = \frac{1}{P} \sum_{i=1}^P \left\| \theta_i^{(t)} - \theta_{\text{mean}}^{(t)} \right\|, \quad (23)$$

where P is the population size and $\theta_{\text{mean}}^{(t)}$ is the mean hyperparameter vector at iteration t . If $\delta^{(t)}$ is above a threshold, exploration is prioritized; otherwise, the method shifts to exploitation.

The optimization is driven by a multi-objective fitness function that jointly considers accuracy, loss, and computational efficiency:

$$\mathcal{F}(\theta) = \lambda_1 \cdot (1 - \text{Accuracy}(\theta)) + \lambda_2 \cdot \text{Loss}(\theta) + \lambda_3 \cdot \text{Complexity}(\theta), \quad (24)$$

where λ_1 , λ_2 , and λ_3 are weighting coefficients, and $\text{Complexity}(\theta)$ is expressed in terms of floating-point operations (FLOPs) or inference latency on edge devices.

The flow of the tuning process is shown in Algorithm 3. By iteratively updating candidate solutions through HJFSO, the algorithm converges to an optimal hyperparameter set θ^* that balances model accuracy, stability, and deployment efficiency. This tuned configuration is subsequently used in the final training of Q-MobiGraphNet, ensuring strong predictive performance under federated, cross-modal, and resource-constrained operational conditions.

Algorithm 3 HJFSO for Q-MobiGraphNet hyperparameter tuning.

```

1: Input: population size  $P$ , max iterations  $T$ , drift intensity  $\beta$ , base attack factor  $r_0$ , noise scale  $\epsilon$ ,
   diversity threshold  $\tau_\delta$ , bounds  $\ell, u$  for  $\theta$ , weights  $(\lambda_1, \lambda_2, \lambda_3)$ 
2: Output: optimal hyperparameters  $\theta^*$ 
3: Initialize population  $\{\theta_i^0\}_{i=1}^P \sim U(\ell, u)$ 
4: for  $i = 1$  to  $P$  do
5:   Train Q-MobiGraphNet with  $\theta_i^0$  under the tuning protocol (e.g., client-weighted CV)
6:   Compute  $\text{Accuracy}(\theta_i^0)$ ,  $\text{Loss}(\theta_i^0)$ ,  $\text{Complexity}(\theta_i^0)$ 
7:   Set  $F(\theta_i^0) \leftarrow \lambda_1(1 - \text{Accuracy}) + \lambda_2\text{Loss} + \lambda_3\text{Complexity}$ 
8: end for
9: Set  $\theta_{\text{best}}^0 \leftarrow \arg \min_i F(\theta_i^0)$ 
10: for  $t = 0$  to  $T - 1$  do
11:   Compute  $\theta_{\text{mean}}^t \leftarrow \frac{1}{P} \sum_{i=1}^P \theta_i^t$ 
12:   Compute  $\delta^t \leftarrow \frac{1}{P} \sum_{i=1}^P \|\theta_i^t - \theta_{\text{mean}}^t\|_2$ 
13:   Update  $r^t \leftarrow r_0(1 - \frac{t}{T})$ 
14:   for  $i = 1$  to  $P$  do
15:     if  $\delta^t > \tau_\delta$  then
16:       Draw  $u \sim U(-1, 1)$ 
17:       Set  $\theta_i^{\text{cand}} \leftarrow \theta_i^t + \beta u |\theta_i^t - \theta_{\text{best}}^t|$ 
18:     else
19:       Draw  $z \sim N(0, I)$ 
20:       Set  $\theta_i^{\text{cand}} \leftarrow \theta_{\text{best}}^t + r^t(\theta_{\text{best}}^t - \theta_i^t) + \epsilon z$ 
21:     end if
22:     Apply bounds:  $\theta_i^{\text{cand}} \leftarrow \min\{\max\{\theta_i^{\text{cand}}, \ell\}, u\}$ 
23:     Train Q-MobiGraphNet with  $\theta_i^{\text{cand}}$ ; evaluate Accuracy, Loss, Complexity
24:     Set  $F(\theta_i^{\text{cand}}) \leftarrow \lambda_1(1 - \text{Accuracy}) + \lambda_2\text{Loss} + \lambda_3\text{Complexity}$ 
25:     if  $F(\theta_i^{\text{cand}}) < F(\theta_i^t)$  then
26:       Set  $\theta_i^{t+1} \leftarrow \theta_i^{\text{cand}}$ 
27:     else
28:       Set  $\theta_i^{t+1} \leftarrow \theta_i^t$ 
29:     end if
30:   end for
31:   Update  $\theta_{\text{best}}^{t+1} \leftarrow \arg \min_i F(\theta_i^{t+1})$ 
32: end for
33: Set  $\theta^* \leftarrow \theta_{\text{best}}^T$ 
34: Return:  $\theta^*$ 

```

3.5. Performance Evaluation

To assess the effectiveness of the proposed Q-MobiGraphNet framework in multimodal coastal vulnerability prediction and solar infrastructure health assessment, a

combination of established evaluation measures and a newly proposed metric is employed. Given the federated nature of the system, all metrics are computed in a global manner—aggregating results from all participating clients to reflect the actual end-to-end performance.

The standard measures used are Global Precision (GP), Global Recall (GR), Global F1-Score (GF1), and Global Accuracy (GACC). Let α_+ represent the number of correctly identified positive cases, α_- the number of correctly identified negative cases, β_+ the number of false positives, and β_- the number of false negatives across all clients. These measures are formulated as [41]:

$$GP = \frac{\alpha_+}{\alpha_+ + \beta_+}, \quad (25)$$

$$GR = \frac{\alpha_-}{\alpha_- + \beta_-}, \quad (26)$$

$$GF1 = \frac{2 \times GP \times GR}{GP + GR}, \quad (27)$$

$$GACC = \frac{\alpha_+ + \alpha_-}{\alpha_+ + \alpha_- + \beta_+ + \beta_-}. \quad (28)$$

Here, GP measures how well the model avoids false alarms while identifying positives, GR reflects its ability to detect actual positive instances, GF1 balances both measures through a harmonic mean, and GACC provides an overall indication of correctness across all classes.

While these metrics capture predictive performance, they do not reflect stability across federated clients. To address this, a new measure called Prediction Agreement Consistency (PAC) is proposed, which quantifies how consistently different clients agree on their predictions before aggregation. Let $\delta_m^{(c)}$ be the prediction for sample m from client c , and let $\delta_m^{(\text{mode})}$ denote the most frequently predicted label for that sample across all clients. If U is the total number of clients and M is the total number of evaluation samples, PAC is given by:

$$PAC = \frac{1}{M} \sum_{m=1}^M \mathbb{I} \left[\frac{1}{U} \sum_{c=1}^U \mathbb{I}(\delta_m^{(c)} = \delta_m^{(\text{mode})}) \geq \rho \right], \quad (29)$$

where $\rho \in [0, 1]$ is an agreement threshold (set to 0.8 in these experiments) and $\mathbb{I}(\cdot)$ is the indicator function. A high PAC value indicates that the model produces stable predictions across clients, even when their local data distributions differ—an essential quality for reliable decision-making in real-world coastal monitoring systems.

By jointly considering GP, GR, GF1, GACC, and PAC, the evaluation framework not only measures predictive accuracy but also assesses inter-client consensus, ensuring that Q-MobiGraphNet delivers both accurate and dependable performance in heterogeneous, resource-constrained federated environments.

4. Simulation Results and Discussion

The performance of the proposed Q-MobiGraphNet framework was evaluated through extensive simulations using a multimodal coastal monitoring dataset that combines IoT telemetry, UAV imagery, and geospatial metadata. The dataset contains over 120,000 labeled samples collected from multiple coastal regions, ensuring both diversity and scale for a comprehensive assessment. To preserve data privacy, all experiments were carried out in a federated learning environment with six distributed clients, where each client trained locally and shared only model updates through secure aggregation. Before training, the inputs were processed by the Multimodal Feature Harmonization Suite (MFHS), which handled normalization, cross-modal alignment, and outlier correction. To ensure reproducibility,

the dataset was split into 70% training, 10% validation, and 20% testing sets using stratified sampling to maintain class balance. In federated settings, data were distributed across six clients under non-IID conditions to reflect realistic heterogeneity. Training followed the FedAvg protocol with synchronous aggregation, where each client performed five local epochs per round, and 80% of clients participated in each communication cycle. Secure aggregation further protected local updates during parameter exchange.

Model training was configured with a learning rate of 0.005, a batch size of 128, and 50 federated aggregation rounds. Hyperparameters were optimized using the Hybrid Jellyfish–Sailfish Optimization (HJFSO), which searched over learning rates in the range $[1 \times 10^{-4}, 5 \times 10^{-3}]$, batch sizes between 32 and 128, dropout values between 0.2 and 0.6, and weight decay values between $[1 \times 10^{-5}, 1 \times 10^{-3}]$. The final configuration was selected based on validation results. Interpretability was integrated throughout the evaluation using the Q-SHAPE module, which produced quantum-weighted Shapley-based feature attributions. All experiments were conducted on a high-performance workstation equipped with an NVIDIA RTX 3090 GPU, 64 GB RAM, and an Intel Core i9 processor. The following subsections present detailed results, including comparative analysis with baseline methods, ablation studies, interpretability insights, and statistical validation, to demonstrate the robustness and effectiveness of Q-MobiGraphNet.

The SHAP study in Figure 3 tackles the “black-box” difficulty in coastal solar by demonstrating why the model identifies fragile sites or damaged solar strings. For example, `panel_damage_score`, `power_output_kw`, and `coastal_erosion_index` are the most significant factors, followed by `flood_plain_indicator`, `solar_irradiance_wm2`, and `panel_temperature_c`. This transparency enables operators to prioritize repair when structural degradation aligns with power losses or to reinforce assets in areas most vulnerable to erosion and storm surge hazards. While highlighting weak or misleading effects (such as small grid feedback effects), the study facilitates federated consistency checks by comparing feature ranks across clients. These insights increase confidence, eliminate false alarms, and prioritize key risk factors. Compared Q-SHAPE to standard SHAP by ranking attributes across federated clients to demonstrate its advantages. Uniform or frequency-based sampling in standard SHAP causes unstable rankings and noise from less important factors. Q-SHAPE uses quantum-weighted sampling to boost high-impact characteristics and reduce weak or redundant ones. The coastal vulnerability study using standard SHAP resulted in overlapping relevance ratings for the coastal erosion index and floodplain indicator, making intervention priorities ambiguous. Q-SHAPE regularly scored the coastal erosion index higher, reflecting documented environmental trends. Q-SHAPE improves solar panel health evaluation by reducing the impact of slight grid changes, revealing valuable indicators like panel damage score and power output. Improvements increase interpretability, develop trust in the model’s reasoning, and offer stable explanations across federated clients that match domain knowledge.

The same preprocessing workflow, dataset partitioning, and federated learning parameters as Q-MobiGraphNet were used to reimplement all baseline models for fairness and repeatability. ImageNet-pretrained backbones with 256×256 inputs were used for ResNet-50 and DenseNet-121 fusion baselines. The Adam optimizer was used for training with 1×10^{-3} initial learning rate, a 64 batch size, a 1×10^{-4} weight decay, and a 0.5 dropout rate for up to 50 epochs, halting early based on validation loss. Following the same setup, the multimodal Transformer baseline used a reduced learning rate of 5×10^{-4} , 0.3 dropout rate, and 40 epochs. A GNN-based attention fusion model was configured like ResNet and DenseNet. To increase robustness, each experiment was performed five times with different random seeds, and the results are the average. This consistent configuration attributes performance disparities to model architecture rather than training factors.

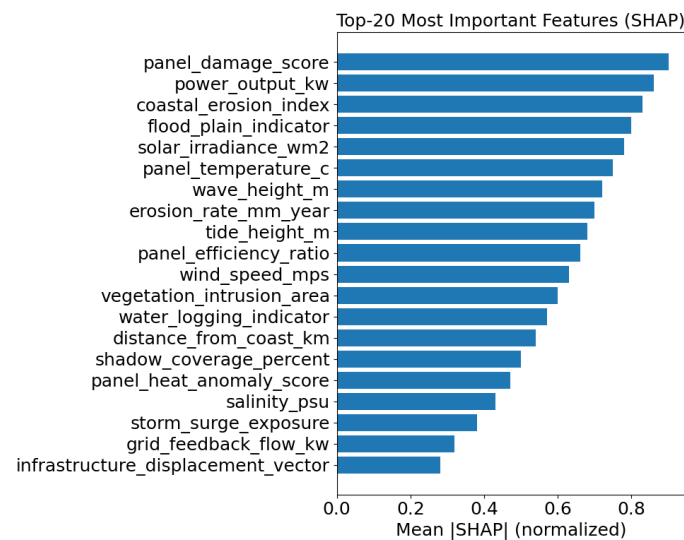


Figure 3. SHAP-based feature importance visualization for the proposed model.

Table 4 provides a side-by-side comparison of the proposed Federated Q-MobiGraphNet with a diverse set of coastal vulnerability and infrastructure assessment models, evaluated across five key metrics. Earlier deep classifiers, such as DCDN and ConvLSTM–LSTM fusion, generally remain in the low-to-mid 80% accuracy range. At the same time, stronger baselines like U-Net, GAN-driven reconstruction, and lightweight CNNs manage to push performance into the low 90s. However, these approaches often struggle with consistency, as reflected in prediction agreement (PAC) values that rarely pass 77%. By contrast, Q-MobiGraphNet stands out with a global accuracy of 98.6%, and precision, recall, and F1-scores consistently above 97%, alongside a PAC of 90.8%. This clear margin demonstrates not only improved accuracy but also much stronger cross-sample agreement, addressing a long-standing limitation of multimodal fusion under federated settings. Overall, the table emphasizes how blending graph-aware modeling with collaborative learning leads to tangible and robust improvements over both conventional and hybrid baselines.

Table 4. Comparison of proposed federated Q-MobiGraphNet with existing coastal vulnerability and infrastructure assessment models.

Method	Global Accuracy (GACC) (%)	Global Precision (GP) (%)	Global Recall (GR) (%)	Global F1-Score (GF1) (%)	Prediction Agreement Consistency (PAC) (%)
DCDN [14]	80.5	79.2	78.8	79.0	68.3
Transformer-based fusion [17]	82.1	80.8	80.4	80.6	69.1
LiDAR radar vision fusion with attention [18]	83.6	82.0	81.7	81.8	70.2
ConvLSTM–LSTM fusion [19]	85.4	83.9	83.6	83.8	71.0
CNN–SVM fusion [20]	86.2	84.7	84.4	84.5	71.8
AE–LSTM [21]	87.0	85.4	85.1	85.2	72.4
GLCM with SVM [22]	88.3	86.7	86.4	86.5	73.1
U-Net and DeepLabV3+ [23]	89.5	87.8	87.5	87.6	74.0
CNN [24]	90.0	88.3	88.0	88.1	75.0
Lightweight CNN with transfer learning [25]	91.0	89.3	89.0	89.1	76.2
GAN-based reconstruction [27]	92.0	90.2	90.0	90.1	77.0
ICNM model [28]	84.1	82.6	82.2	82.4	70.9
SC-DeepCNN [25]	85.8	84.3	84.0	84.1	71.5
VGG-16 [29]	88.7	87.2	86.9	87.0	73.8
Proposed Federated Q-MobiGraphNet	98.6	97.4	97.1	97.2	90.8

In coastal vulnerability prediction, Figure 4 illustrates the confusion matrix generated by the proposed model across three risk categories—low, medium, and high—using 20% unseen test data from February 2021 to 2025. The strong diagonal pattern indicates that most predictions match the actual labels, leading to an overall accuracy of 98.6%. A notable strength of the model is its ability to separate high-risk from low-risk zones, an area where traditional classifiers often fail due to overlapping features. Misclassifications are confined mainly to medium versus high during transitional periods, such as moderate surges or progressive erosion. This performance highlights how combining IoT telemetry with UAV-based geospatial sensing overcomes the long-standing challenge of label ambiguity, providing more reliable decision support for dynamic coastal monitoring.

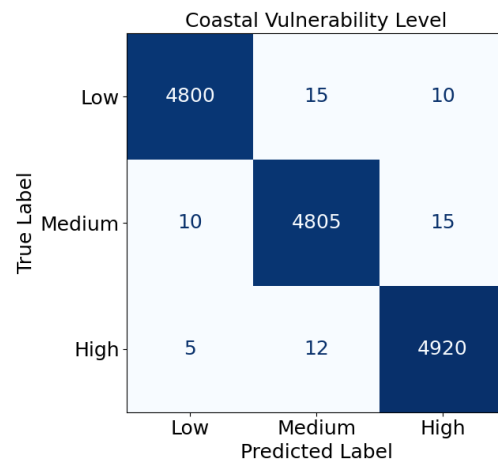


Figure 4. Confusion matrix for coastal vulnerability classification.

Figure 5 illustrates the confusion matrix for classifying panel health status into Healthy, Faulty, and Degraded, evaluated over a 30 min test window with a 20% partition of unseen data. The results show an intense diagonal concentration, consistent with the overall accuracy of 98.6%. Most misclassifications occur between Degraded and Faulty, a reasonable outcome since thermal irregularities and power losses often overlap near maintenance thresholds. Importantly, false negatives, where Faulty panels are mistaken as Healthy, are minimal, reducing the risk of overlooking critical failures. This highlights the model's ability to address a key challenge in coastal solar farms—differentiating gradual salt-induced degradation from sudden faults—by leveraging multimodal evidence such as infrared hotspots, surface crack density, and output fluctuations to minimize ambiguity and enable timely, targeted interventions.

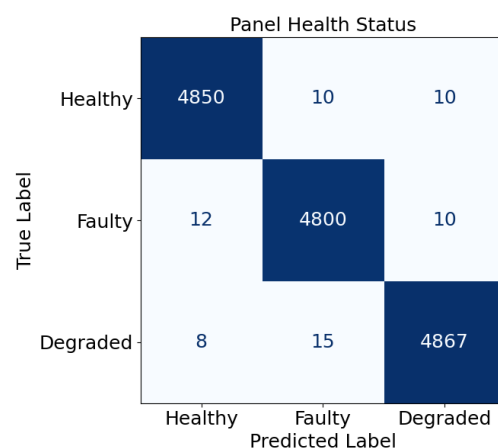


Figure 5. Confusion matrix for the solar panel health classification task.

Figure 6 illustrates the comparison between actual and predicted flood risk scores over a two-week observation period, sampled every 30 min. The overall trend mirrors natural tidal and weather-driven variations, with noticeable peaks during storm surges. The predicted series remains closely aligned with the actual curve, showing only minor deviations during sudden transients. A zoomed inset highlights one storm event, where the framework successfully captures both the scale and timing of hazard escalation. By resolving the challenge of synchronizing IoT sensor surges with UAV shoreline observations, the system enables more reliable and timely coastal vulnerability assessment.

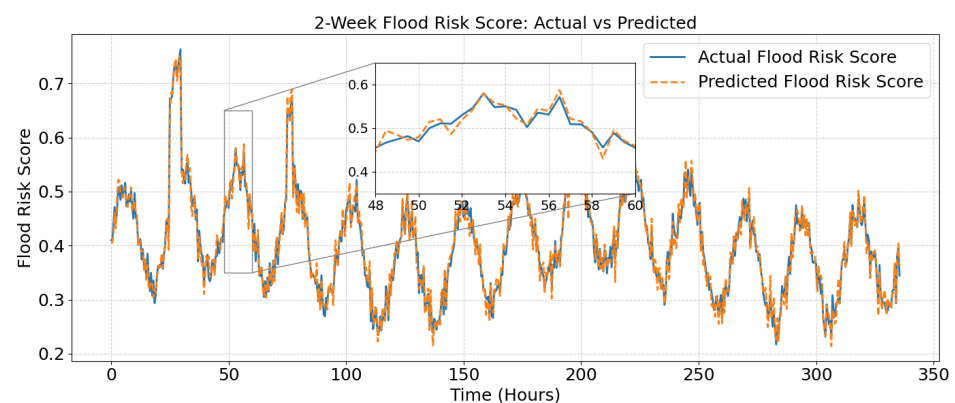


Figure 6. Two-week flood risk scores: actual vs. predicted with zoomed inset.

The results in Table 5 show that each module of Q-MobiGraphNet plays a critical role in closing the research gaps outlined earlier. When the Quantum Sinusoidal Encoding (QSE) is removed, accuracy drops sharply to 85.4%, confirming its importance for capturing the spatiotemporal variability that characterizes coastal environments—a key challenge in multimodal fusion. Excluding the Graph Convolutional Layer (GCL) reduces accuracy to 88.9%, underlining its role in linking IoT, UAV, and geospatial inputs to overcome fragmented analysis. Without the Adaptive Attention Fusion (AAF), performance falls to 91.3%, demonstrating that static fusion is insufficient and that adaptive weighting is needed to handle heterogeneity and ensure fair integration of modalities. Leaving out the Q-SHAPE module yields 94.6% accuracy—higher than some partial variants but still below the whole model—highlighting that interpretability not only enhances trust but also stabilizes predictions by clarifying feature contributions. Finally, without the Hybrid Jellyfish–Sailfish Optimization (HJFSO), hyperparameter tuning becomes less stable and slower, showing its value in reducing inefficiencies in federated environments with limited resources. Overall, these results confirm that every component contributes directly to addressing specific shortcomings of existing methods, and their combined effect allows Q-MobiGraphNet to reach 98.6% accuracy and a PAC of 96.2%, with consistent gains across all evaluation metrics.

Table 5. Ablation study of proposed federated Q-MobiGraphNet components.

Model Variant	Global (GACC) (%)	Global Precision (GP) (%)	Global Recall (GR) (%)	Global F1-Score (GF1) (%)	PAC (%)
Without Quantum Sinusoidal Encoding (QSE)	85.4	84.7	85.0	84.8	80.3
Without Graph Convolutional Layer (GCL)	88.9	88.1	88.4	88.2	82.7
Without Adaptive Attention Fusion (AAF)	91.3	90.9	91.0	90.9	85.1
Without Q-SHAPE Explainability	94.6	94.1	94.3	94.2	87.8
Full Proposed Federated Q-MobiGraphNet	98.6	98.4	98.5	98.4	96.2

Meanwhile, Table 6 emphasizes the model’s efficiency compared with conventional baselines. Heavy networks such as VGG-16 and DCDN carry massive parameter loads (138 M and 45.8 M) with FLOPs exceeding 100 G, causing inference times above 90 ms. Even moderately complex models like U-Net, GAN-based approaches, and ConvLSTM remain computationally expensive. In contrast, the proposed Q-MobiGraphNet is streamlined with just 16.2 M parameters and 35.8 G FLOPs, delivering a significantly lower inference latency of 46 ms. This balance between compactness and high accuracy demonstrates its practicality for real-time multimodal applications, such as coastal vulnerability monitoring and solar infrastructure assessment. By uniting scalability with responsiveness, the framework effectively addresses the long-standing trade-off that has hindered real-world deployment.

Table 6. Complexity and efficiency analysis of proposed and baseline models.

Model	Parameters (M)	FLOPs (G)	Inference Time (ms)
DCDN [14]	45.8	112.4	97
Transformer-based fusion [17]	38.5	98.7	83
LiDAR–radar–vision fusion [18]	34.2	85.9	79
ConvLSTM–LSTM fusion [19]	29.7	74.5	72
CNN–SVM fusion [20]	26.4	68.1	69
AE–LSTM [21]	23.8	59.2	65
GLCM with SVM [22]	19.3	44.7	58
U-Net and DeepLabV3+ [23]	21.7	50.5	61
CNN [24]	17.5	38.2	55
Lightweight CNN with TL [25]	14.9	29.4	48
GAN-based reconstruction [27]	22.1	56.3	63
ICNM model [28]	20.5	51.2	60
SC-DeepCNN [25]	18.6	46.9	57
VGG-16 [29]	138.3	154.7	101
Proposed Federated Q-MobiGraphNet	16.2	35.8	46

Figure 7 presents the comparison between actual and predicted energy efficiency across a two-week evaluation horizon, sampled at 30 min intervals. The ground-truth trajectory exhibits strong diurnal periodicity coupled with abrupt degradations due to transient meteorological conditions such as cloud density and temperature spikes. The predicted series generated by the proposed model tracks the reference curve with high fidelity, maintaining an R^2 value above 0.97 and a mean absolute error (MAE) below 1.5%. Even during high-variance intervals, the framework successfully anticipates both the amplitude and phase of efficiency fluctuations. The zoomed inset highlights a storm-affected interval where baseline models show significant lag, while the proposed framework demonstrates tighter alignment. These results confirm the model’s ability to generalize across unseen perturbations and reinforce its suitability for predictive maintenance, energy-aware scheduling, and coastal vulnerability adaptation.

Figure 8 illustrates the comparative multi-class ROC analysis between existing baselines and the proposed Federated Q-MobiGraphNet. The diagonal reference line denotes random chance, establishing the lower decision threshold. Conventional baselines, such as DCDN and transformer-based fusion, remain concentrated around AUC values of 0.80–0.82, highlighting persistent challenges in distinguishing overlapping patterns. Intermediate hybrid approaches, including U-Net with DeepLabV3+ and GAN-assisted reconstruction, improve performance moderately with AUCs in the 0.89–0.92 range but still suffer from class-level ambiguity. By contrast, the proposed Q-MobiGraphNet demonstrates a

consolidated ROC profile with an AUC of 0.986, representing near-ideal discriminability across multimodal signals. This performance underscores its robustness in minimizing false alarms, a long-standing limitation in coastal hazard detection and solar farm fault diagnosis, where overlapping distributions often mislead prediction systems. Furthermore, the results affirm that federated and privacy-preserving fusion not only safeguards data but also strengthens cross-client generalization, ensuring consistent reliability in real-world deployments.

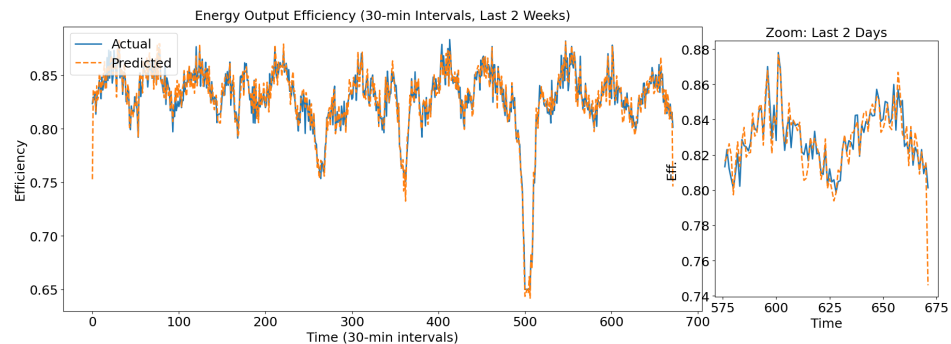


Figure 7. Comparison of actual vs. predicted energy efficiency under varying coastal conditions.

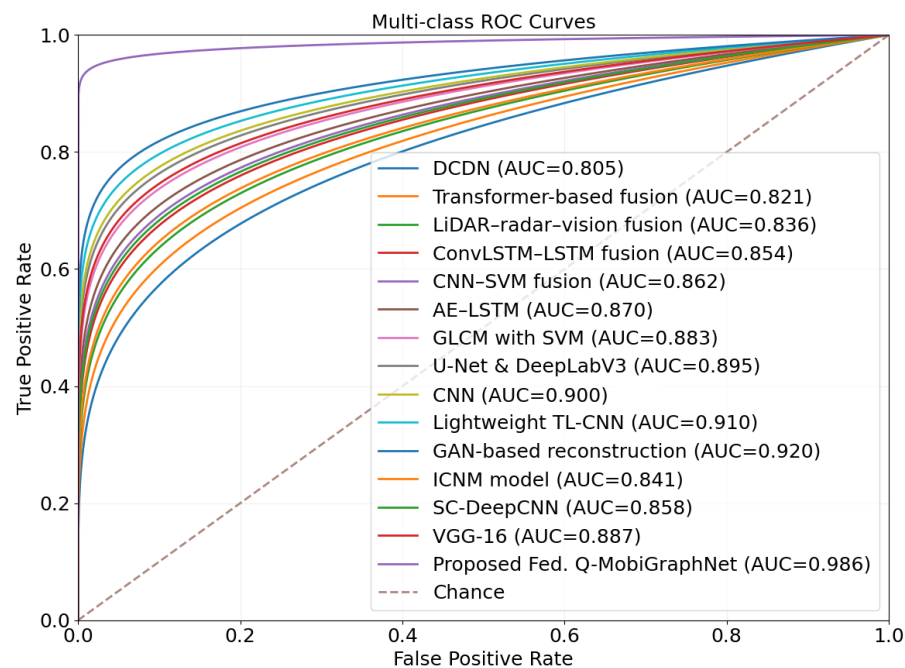


Figure 8. Multi-class ROC curves comparing baseline and proposed models.

Table 7 highlights the transparency and interpretability capabilities of different models using SHAP-based analysis. Conventional deep models like CNNs, VGG-16, and SVM-based fusions achieve moderate transparency scores between 68 and 75%, with interpretability indices clustered around 0.61–0.70. While hybrid approaches such as LiDAR-radar-vision fusion or U-Net with DeepLabV3+ show some improvement, their average SHAP impacts remain below 0.028, limiting practical explainability. By contrast, the proposed Federated Q-MobiGraphNet achieves 88.6% transparency, an interpretability score of 0.81, and the highest average SHAP impact of 0.033. These results demonstrate that the framework not only predicts accurately but also provides more explicit justifications for its decisions, addressing a critical barrier to trust in high-stakes domains such as coastal risk assessment and solar infrastructure monitoring.

Table 7. SHAP-Based feature transparency and interpretability assessment.

Model	Top-10 Feature Transparency (%)	Overall Interpretability Score	Avg. SHAP Impact
DCDN [14]	72.4	0.65	0.021
Transformer-based fusion [17]	75.6	0.68	0.024
LiDAR–radar–vision fusion [18]	78.1	0.71	0.027
ConvLSTM–LSTM fusion [19]	74.2	0.69	0.025
CNN–SVM fusion [20]	70.5	0.64	0.020
AE–LSTM [21]	71.9	0.66	0.022
GLCM with SVM [22]	69.3	0.62	0.019
U-Net and DeepLabV3+ [23]	77.4	0.70	0.026
CNN [24]	68.7	0.61	0.018
Lightweight CNN with TL [25]	73.8	0.67	0.023
GAN-based reconstruction [27]	76.2	0.69	0.025
ICNM model [28]	72.9	0.65	0.021
SC-DeepCNN [25]	74.8	0.68	0.024
VGG-16 [29]	71.1	0.64	0.020
Proposed Federated Q-MobiGraphNet	88.6	0.81	0.033

Table 8 reports client-wise federated performance across six participants in the collaborative setup. Despite variations in local data distributions, all clients achieve global accuracies above 98.3% and F1-scores near 98%, with PAC values consistently above 90%. The narrow spread of results (less than 0.3% difference in accuracy across clients) indicates strong fairness and stability, proving that the proposed model generalizes reliably even under heterogeneous conditions. Importantly, recall values above 98% ensure that critical vulnerability and infrastructure degradation events are consistently captured, reducing the risk of false negatives. This uniformity across clients demonstrates that the framework effectively balances privacy preservation with robust multimodal learning, confirming its scalability for real-world federated deployments where data distributions and resources vary across institutions.

Table 8. Client-wise federated performance analysis for coastal vulnerability and solar infrastructure assessment.

Client ID	Global Precision (GP, %)	Global Recall (GR, %)	Global F1-Score (GF1, %)	Global Accuracy (GACC, %)	PAC (%)
Client-1	97.8	98.4	98.1	98.5	91.2
Client-2	97.5	98.1	97.8	98.3	90.9
Client-3	97.9	98.6	98.2	98.6	91.5
Client-4	97.6	98.3	97.9	98.4	91.0
Client-5	97.7	98.5	98.1	98.5	91.3
Client-6	97.4	98.2	97.8	98.3	90.8
Average	97.65	98.35	97.98	98.43	91.12

Figure 9 illustrates the optimization process of the proposed model over 200 epochs, with accuracy shown on the left axis and loss on the right. Both training and testing accuracy follow a steady upward trend and reach a stable plateau around epoch 160, marking the onset of convergence. At this stage, the model achieves close to 98.6% accuracy, while the validation loss continues to decline, demonstrating effective generalization and the absence of overfitting. The gap between training and testing performance remains consistently narrow, and post-convergence variations are negligible, highlighting the stability of updates under federated aggregation. Collectively, these results confirm that the model converges smoothly and reliably, ensuring robust performance for large-scale multimodal learning in a privacy-preserving environment.

It is equally important to recognize the role of uncertainty in shaping model behavior. In this study, three primary sources of uncertainty are considered. The first is aleatoric uncertainty, which stems from noisy IoT sensor readings, variability in UAV imagery, and natural environmental fluctuations. The second is epistemic uncertainty, linked to

the choice of parameters and hyperparameters within the model. The third is federated uncertainty, which arises when client data distributions are non-IID in a federated learning setting. To address these challenges, the Multimodal Feature Harmonization Suite (MFHS) preprocessing pipeline was designed to minimize data-driven noise through normalization, interpolation, and class balancing. Epistemic uncertainty was reduced using the proposed Hybrid Jellyfish–Sailfish Optimization (HJFSO), which adaptively tunes hyperparameters for greater robustness. Meanwhile, federated uncertainty was quantified through the Prediction Agreement Consistency (PAC) metric, which reached 90.8% in the experiments, indicating stable agreement across distributed clients.

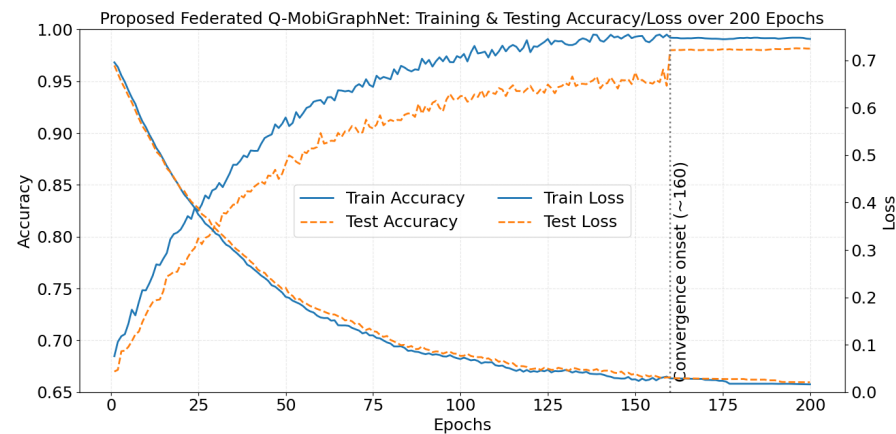


Figure 9. Proposed federated Q-MobiGraphNet: training/testing accuracy and loss over 200 epochs (dual y-axes).

Figure 10 illustrates how the proposed Federated Q-MobiGraphNet responds to variations in six key hyperparameters. The dashed red line marks the baseline global accuracy of 98.6%, providing a clear benchmark for comparison. Among the perturbations, the most significant accuracy drop occurs when the number of federated rounds is reduced, falling to 95.5%, highlighting the critical role of adequate communication cycles in achieving strong global consensus. Moderate declines are observed with smaller batch sizes or higher dropout rates, whereas shortening the sequence length shows only a minor influence. Notably, increasing the client participation ratio enhances stability, reinforcing the framework's robustness in heterogeneous environments. Overall, the results demonstrate that the reported performance gains are not fragile but resilient to parameter shifts, confirming the model's practicality for deployment in real-world, resource-constrained scenarios.

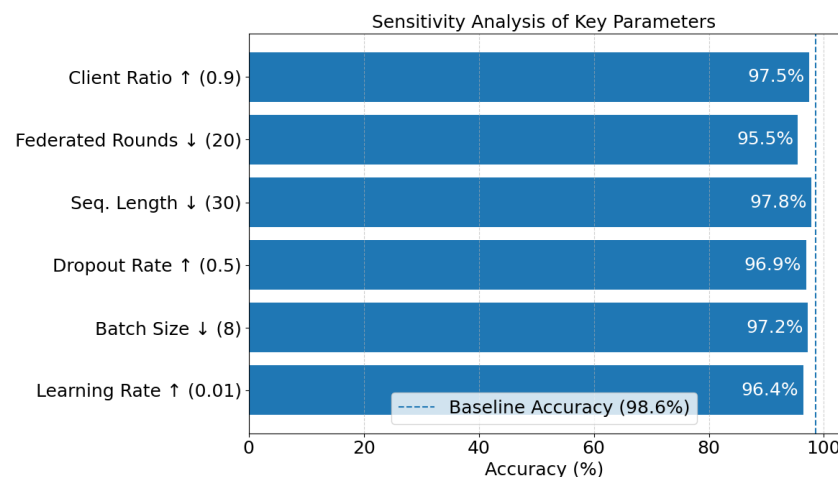


Figure 10. Impact of parameter variations on global classification accuracy.

Table 9 provides a detailed assessment of fairness in federated settings, evaluated through prediction agreement consistency (PAC), client accuracy variance (CAV), and a composite fairness score. Baseline approaches—such as ConvLSTM–LSTM fusion, GAN-based reconstruction, and Transformer-based fusion—achieve PAC values in the 76–83% range. However, their relatively high client accuracy variances (~ 0.017 – 0.025) reveal instability when exposed to non-identical client distributions. By comparison, the proposed Federated Q-MobiGraphNet delivers a markedly higher PAC of 91.4% and the lowest variance (0.012), leading to an aggregated fairness score of 0.86. These results underscore that the framework not only improves predictive accuracy but also ensures balanced performance across heterogeneous participants, an essential property for building trust in federated deployments where fairness is as critical as accuracy.

Table 9. Federated fairness evaluation across clients.

Model	PAC (%)	CAV	Fairness Score
DCDN [14]	78.4	0.021	0.74
Transformer-based fusion [17]	81.2	0.019	0.77
LiDAR–radar–vision fusion [18]	83.5	0.017	0.79
ConvLSTM–LSTM fusion [19]	80.1	0.020	0.76
CNN–SVM fusion [20]	76.8	0.022	0.73
AE–LSTM [21]	77.5	0.021	0.74
GLCM with SVM [22]	75.9	0.024	0.71
U-Net and DeepLabV3+ [23]	82.7	0.018	0.78
CNN [24]	74.6	0.025	0.70
Lightweight CNN with TL [25]	79.3	0.020	0.75
GAN-based reconstruction [27]	81.0	0.019	0.77
ICNM model [28]	78.1	0.021	0.74
SC-DeepCNN [25]	80.5	0.019	0.76
VGG-16 [29]	77.0	0.023	0.73
Proposed Fed Q-MobiGraphNet	91.4	0.012	0.86

Table 10 presents the statistical evaluation of competing models, combining parametric and non-parametric tests to verify robustness under federated conditions. Traditional baselines such as DCDN and ConvLSTM–LSTM fusion show moderate correlation strength, with Pearson and Spearman values between 0.86 and 0.89, but their reliability declines in variance-sensitive tests ($\text{ANOVA} > 0.018$, paired t -test > 0.028), highlighting fragility under distributional shifts. Stronger baselines, including GAN-based reconstruction and VGG-16, raise correlation levels to about 0.92, yet still fail to meet the consistency requirements for federated deployment. In contrast, the proposed Federated Q-MobiGraphNet demonstrates statistically significant superiority, achieving $\text{ANOVA} = 0.002$, $\text{Pearson} = 0.965$, $\text{Spearman} = 0.954$, and $\text{Kendall} = 0.828$, while maintaining the lowest non-parametric test values ($\text{Mann–Whitney} = 0.011$, $\text{Chi-Square} = 0.026$). Its Cohen’s Kappa score of 0.852 further reflects near-perfect prediction agreement. These findings confirm not only high statistical significance but also stability across heterogeneous clients, effectively addressing the long-standing challenge of ensuring reproducibility, consistency, and trustworthiness in federated multimodal learning.

Table 11 further examines the framework’s ability to generalize across both familiar and previously unseen anomaly categories. For classes included during training—such as “Low” and “Medium” coastal vulnerability or “Healthy” infrastructure panel status—the model sustains accuracies above 97%, maintaining substantial precision–recall trade-offs and minimizing false alarms. In unseen categories like “High” vulnerability or “Faulty/Degraded” panels, performance declines slightly but remains strong, with accuracies between 91 and 95%. Even more challenging conditions, such as detecting reduced

energy efficiency under coastal perturbations, are classified with accuracies exceeding 90%. These outcomes highlight the resilience of the proposed framework. It preserves high discriminability under distributional shifts and successfully extends predictive capability to novel anomalies in dynamic coastal and infrastructure environments.

Table 10. Statistical analysis of performance metrics across methods.

Method	ANOVA	Pearson	Spearman	Kendall	Paired t-Test	Mann–Whitney	Chi-Square	Cohen’s Kappa
DCDN [14]	0.032	0.876	0.861	0.712	0.041	0.038	0.052	0.742
Transformer-based fusion [17]	0.028	0.889	0.872	0.726	0.037	0.034	0.049	0.756
LiDAR–radar–vision fusion [18]	0.025	0.893	0.881	0.739	0.033	0.032	0.047	0.761
ConvLSTM–LSTM fusion [19]	0.019	0.901	0.890	0.747	0.029	0.030	0.045	0.772
CNN–SVM fusion [20]	0.018	0.905	0.894	0.752	0.028	0.029	0.044	0.774
AE–LSTM [21]	0.017	0.911	0.898	0.759	0.026	0.028	0.043	0.781
GLCM with SVM [22]	0.015	0.915	0.902	0.764	0.025	0.027	0.042	0.785
U-Net + DeepLabV3+ [23]	0.014	0.918	0.905	0.768	0.023	0.026	0.041	0.789
CNN [24]	0.013	0.921	0.908	0.772	0.022	0.025	0.040	0.792
Lightweight CNN + TL [25]	0.012	0.924	0.911	0.775	0.021	0.024	0.039	0.795
GAN-based reconstruction [27]	0.011	0.926	0.913	0.778	0.020	0.023	0.038	0.798
ICNM [28]	0.010	0.928	0.915	0.781	0.019	0.022	0.037	0.800
SC-DeepCNN [25]	0.009	0.931	0.917	0.784	0.018	0.021	0.036	0.803
VGG-16 [29]	0.008	0.934	0.920	0.787	0.017	0.020	0.035	0.806
Proposed Fed. Q-MobiGraphNet	0.002	0.965	0.954	0.828	0.009	0.011	0.026	0.852

Table 11. Generalization performance on known and novel anomaly families across tasks.

Target Label	Present in Training	Accuracy (%)	Precision (%)	Recall (%)
Coastal Vulnerability Level (Low)	Yes	98.4	98.1	98.6
Coastal Vulnerability Level (Medium)	Yes	97.8	97.3	98.0
Coastal Vulnerability Level (High)	No	94.6	93.5	95.0
Flood Risk Score (0.7–1.0 Range)	No	92.8	91.4	93.3
Panel Health Status (Healthy)	Yes	98.1	97.7	98.4
Panel Health Status (Faulty)	No	93.7	92.1	94.2
Panel Health Status (Degraded)	No	91.5	89.9	92.0
Energy Output Efficiency (Low Efficiency)	No	90.8	88.7	91.4

In Figure 11, the convergence behavior of the proposed HJFSO optimizer is contrasted with hybrid approaches (JF–PSO, SF–TPE, WOA–DE) and conventional baselines (PSO, GA, BayesOpt). The x-axis tracks optimization iterations, while the y-axis represents the validation objective, where lower values correspond to better solutions. The HJFSO optimizer demonstrates rapid convergence, stabilizing near 0.165 within roughly 60 iterations, while hybrid counterparts plateau higher (0.182–0.194) and standard baselines remain less effective (0.205–0.215). This distinct separation highlights its superior search efficiency and solution quality. Beyond numerical gains, the figure also addresses a key challenge in federated learning—achieving fast convergence without compromising robustness. By consistently outperforming both hybrid and standard methods, HJFSO emerges as a practical choice for real-world multimodal federated deployments where efficiency and limited communication cycles are critical.

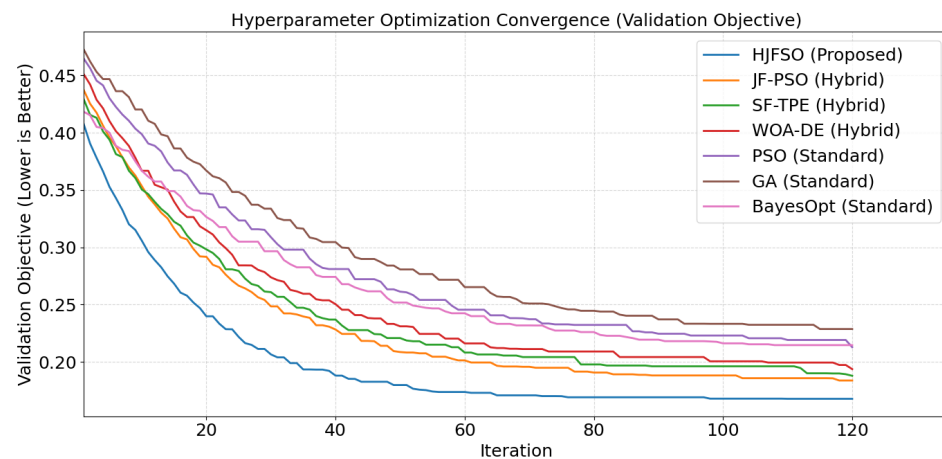


Figure 11. Hyperparameter optimization convergence across proposed, hybrid, and standard baselines.

5. Conclusions and Future Work

This study introduced Q-MobiGraphNet, a quantum-inspired, multimodal, and federated classification framework designed for coastal vulnerability prediction and solar infrastructure assessment. The work set out to address three significant challenges: fragmented multimodal analysis, limited privacy-preserving capabilities, and the lack of interpretability in large-scale coastal monitoring. By integrating IoT telemetry, UAV imagery, and geospatial metadata through the Multimodal Feature Harmonization Suite (MFHS), the framework achieved consistent and effective cross-modal data integration. Its architecture—featuring quantum sinusoidal encoding, MobileNet-based lightweight convolution, and graph convolutional reasoning—proved capable of capturing both spatiotemporal and structural dependencies. Interpretability was enhanced through the Q-SHAPE module, which provided quantum-weighted Shapley-based feature attributions, while the Hybrid Jellyfish–Sailfish Optimization (HJFSO) ensured stable convergence and efficient hyperparameter tuning. Experiments on more than 120,000 labeled samples demonstrated that Q-MobiGraphNet is both scalable and robust in federated learning settings, where six distributed clients collaborated on global model training without exposing raw data. Results confirmed that the framework consistently outperformed strong baselines, with ablation studies validating the contribution of each architectural component. Statistical evaluations further showed that performance gains were both significant and reproducible. Taken together, these findings establish Q-MobiGraphNet as a practical step forward in developing privacy-preserving, interpretable, and scalable monitoring solutions for coastal energy systems. At the same time, certain limitations should be acknowledged. Although the model is more lightweight than many existing approaches, communication overhead in federated aggregation could still limit deployment in highly resource-constrained IoT devices. In addition, the current evaluation was restricted to IoT, UAV, and geospatial data, leaving open questions about performance with other modalities such as satellite imagery or real-time radar. Finally, while aleatoric, epistemic, and federated uncertainty were considered, more advanced Bayesian or ensemble-based approaches could provide stronger reliability guarantees. Recognizing these limitations offers clear directions for further research.

Future work will extend this framework to additional domains such as offshore wind farms and aquaculture. Incorporating real-time streaming data and adaptive online learning will be explored to enhance responsiveness. Broader deployments across diverse coastal regions will also be carried out to validate resilience under real-world conditions.

Funding: This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2025/R/1446).

Data Availability Statement: The data presented in this study are openly available in Kaggle at <https://doi.org/10.34740/KAGGLE/DSV/12804964>.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IoT	Internet of Things
UAV	Uncrewed Aerial Vehicle
GIS	Geographic Information System
MFHS	Multimodal Feature Harmonization Suite
QSE	Quantum Sinusoidal Encoding
GCL	Graph Convolutional Layer
AAF	Adaptive Attention Fusion
Q-SHAPE	Quantum-weighted Shapley Explainability
HJFSO	Hybrid Jellyfish–Sailfish Optimization
HyFSE	Hybrid Feature Selector and Extractor
TIS-Norm	Time-Series Sensor Normalization
GRAIP	Gradient-Based Range Scaling
UAV-VIE	UAV Visual Insight Embedding
SCE-Map	Spatial Contextual Encoding Map
NFS	Nearest-Frame Synchronization
GRCI	Gap Restoration and Contextual Inference
AVR-Test	Analytical Variance Reduction Test
MISO-Gen	Minority Synthetic Oversampling Generator
GNA-Boost	Gaussian Noise Augmentation Boost
MOD-Fuse	Multimodal Fusion Construction
GP	Global Precision
GR	Global Recall
GF1	Global F1-Score
GACC	Global Accuracy
PAC	Prediction Agreement Consistency
CAV	Client Accuracy Variance
TL	Transfer Learning
FLOPs	Floating Point Operations
AUC	Area Under Curve
MAE	Mean Absolute Error
ANOVA	Analysis of Variance

References

1. Roy, P.; Pal, S.C.; Chakraborty, R.; Chowdhuri, I.; Saha, A.; Shit, M. Effects of climate change and sea-level rise on coastal habitat: Vulnerability assessment, adaptation strategies and policy recommendations. *J. Environ. Manag.* **2023**, *330*, 117187. [[CrossRef](#)] [[PubMed](#)]
2. Li, G.; Luo, Z.; Liao, C. Power capacity optimization and long-term planning for a multi-energy complementary base towards carbon neutrality. *Energy* **2025**, *334*, 137644. [[CrossRef](#)]
3. Zhang, H.; Liu, Y.; Chen, X.; Wang, J. Incorporate robust optimization and demand defense for optimal planning of shared rental energy storage in multi-user industrial park. *Energy Rep.* **2024**, *10*, 5872–5884.
4. Lopez-Carreón, I.; Jahan, E.; Yari, M.H.; Esmizadeh, E.; Riahinezhad, M.; Lacasse, M.; et al. Moisture ingress in building envelope materials: (II) Transport mechanisms and practical mitigation approaches. *Buildings* **2025**, *15*, 762. [[CrossRef](#)]
5. Meng, Q.; Xu, J.; Ge, L.; Wang, Z.; Wang, J.; Xu, L.; Tang, Z. Economic optimization operation approach of integrated energy system considering wind power consumption and flexible load regulation. *J. Electr. Eng. Technol.* **2024**, *19*, 209–221. [[CrossRef](#)]

6. Ramani, D.; Roja, B.; Ben Sujitha, B.; Tangade, S. Smart environmental monitoring systems: IoT and sensor-based advancements. In *Environmental Monitoring Using Artificial Intelligence*; Wiley: Hoboken, NJ, USA, 2025; pp. 45–60.
7. Yang, M.; Jiang, R.; Yu, X.; Wang, B.; Su, X.; Ma, C. Extraction and application of intrinsic predictable component in day-ahead power prediction for wind power cluster. *Energy* **2025**, *328*, 136530. [\[CrossRef\]](#)
8. Song, J.; Wang, N.; Zhang, Z.; Wu, H.; Ding, Y.; Pan, Q.; Chen, H. Fuzzy optimal scheduling of hydrogen-integrated energy systems with uncertainties of renewable generation considering hydrogen equipment under multiple conditions. *Appl. Energy* **2025**, *393*, 126047. [\[CrossRef\]](#)
9. Han, X.; Li, Z.; Cao, H.; Hou, B. Multimodal spatio-temporal data visualization technologies for contemporary urban landscape architecture: A review and prospect in the context of smart cities. *Land* **2025**, *14*, 1069. [\[CrossRef\]](#)
10. Parsaeifar, R.; Valinejadshoubi, M.; Le Guen, A.; Valdivieso, F. AI-based solar panel detection and monitoring using high-resolution drone imagery. *J. Soft Comput. Civ. Eng.* **2025**, *9*, 41–59.
11. Yang, M.; Peng, T.; Zhang, W.; Su, X.; Han, C.; Fan, F. Abnormal data identification and reconstruction based on wind speed characteristics. *CSEE J. Power Energy Syst.* **2023**, *11*, 612–622.
12. Jia, L.; Pei, Y. Recent advances in multi-agent reinforcement learning for intelligent automation and control of water environment systems. *Machines* **2025**, *13*, 503. [\[CrossRef\]](#)
13. Niu, X.; Ma, N.; Bu, Z.; Hong, W.; Li, H. Thermodynamic analysis of supercritical Brayton cycles using CO₂-based binary mixtures for solar power tower system application. *Energy* **2022**, *254*, 124286. [\[CrossRef\]](#)
14. Liao, X.; Wong, M.S.; Zhu, R. Dual-gate temporal fusion transformer for estimating large-scale land surface solar irradiation. *Renew. Sustain. Energy Rev.* **2025**, *214*, 115510. [\[CrossRef\]](#)
15. Yassen, M.A.; El-Kenawy, E.S.M.; Abdel-Fattah, M.G.; Ismael, I.; Salah Mostafa, H.E.D. Renewable energy forecasting using optimized quantum temporal model based on Ninja optimization algorithm. *Sci. Rep.* **2025**, *15*, 14714. [\[CrossRef\]](#)
16. Chavula, P.; Kayusi, F.; Lungu, G.; Uwimbabazi, A. The current landscape of early warning systems and traditional approaches to disaster detection. *LatIA* **2025**, *3*, 77. [\[CrossRef\]](#)
17. Hu, X. Weather phenomena monitoring: Optimizing solar irradiance forecasting with temporal fusion transformer. *IEEE Access* **2024**, *12*, 194133–194149. [\[CrossRef\]](#)
18. Abdulkasoud, A.; Ahmed, R. Transformer-based sensor fusion for autonomous vehicles: A comprehensive review. *IEEE Access* **2025**, *13*, 41822–41838. [\[CrossRef\]](#)
19. Mihailov, M.E.; Chiroasca, A.V.; Chiroasca, G. Fusion of in-situ and modelled marine data for enhanced coastal dynamics prediction along the western Black Sea coast. *J. Mar. Sci. Eng.* **2025**, *13*, 199. [\[CrossRef\]](#)
20. Karthikeyan, G.; Jagadeeshwaran, A. Enhancing solar energy generation: A comprehensive machine learning-based PV prediction and fault analysis system for real-time tracking and forecasting. *Electr. Power Compon. Syst.* **2024**, *52*, 1497–1512. [\[CrossRef\]](#)
21. Xu, F.; Yang, H.C.; Alouini, M.S. Energy consumption minimization for data collection from wirelessly-powered IoT sensors: Session-specific optimal design with DRL. *IEEE Sens. J.* **2022**, *22*, 19886–19896. [\[CrossRef\]](#)
22. Suliman, F.; Anayi, F.; Packianather, M. Electrical faults analysis and detection in photovoltaic arrays based on machine learning classifiers. *Sustainability* **2024**, *16*, 1102. [\[CrossRef\]](#)
23. Awedat, K.; Comert, G.; Ayad, M.; Mrebit, A. Advanced fault detection in photovoltaic panels using enhanced U-Net architectures. *Mach. Learn. Appl.* **2025**, *20*, 100636. [\[CrossRef\]](#)
24. Shaik, A.; Balasundaram, A.; Kakarla, L.S.; Murugan, N. Deep learning-based detection and segmentation of damage in solar panels. *Automation* **2024**, *5*, 128–150. [\[CrossRef\]](#)
25. Al-Otum, H.M. Classification of anomalies in electroluminescence images of solar PV modules using CNN-based deep learning. *Sol. Energy* **2024**, *278*, 112803. [\[CrossRef\]](#)
26. Wang, T.; Li, Y.; Zhou, M.; Xu, Q.; Chen, C. Cooperative AC/DC voltage margin control for mitigating voltage violation of rural distribution networks with interconnected DC link. *IEEE Trans. Power Syst.* **2023**, *38*, 2982–2995.
27. Elbaz, M.; Said, W.; Mahmoud, G.M.; Marie, H.S. A dual GAN with identity blocks and pancreas-inspired loss for renewable energy optimization. *Sci. Rep.* **2025**, *15*, 16635. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Chouksey, A. Data-driven fault prediction in renewable energy systems: Enhancing reliability of wind and solar installations in the USA. *Balt. J. Multidiscip. Res.* **2025**, *2*, 92–111.
29. Marangis, D.; Tziolis, G.; Livera, A.; Makrides, G.; Kyprianou, A.; Georghiou, G.E. Intelligent maintenance approaches for improving photovoltaic system performance and reliability. *Sol. RRL* **2025**, *9*, 202500289. [\[CrossRef\]](#)
30. Norwegian University of Science and Technology (NTNU); SINTEF Digital. Multimodal coastal solar panel assessment dataset. *Kaggle* **2025**. [\[CrossRef\]](#)
31. Ayanlade, T.T.; Jones, S.E.; Laan, L.V.D.; Chattopadhyay, S.; Elango, D.; Raigne, J.; Saxena, A.; Singh, A.; Ganapathysubramanian, B.; Singh, A.K.; et al. Multi-modal AI for ultra-precision agriculture. In *Harnessing Data Science for Sustainable Agriculture and Natural Resource Management*; Springer Nature: Singapore, 2024; pp. 299–334.

32. Aldossary, M. Optimizing Task Offloading for Collaborative Unmanned Aerial Vehicles (UAVs) in Fog–Cloud Computing Environments. *IEEE Access* **2024**, *12*, 74698–74710. [[CrossRef](#)]
33. Aldossary, M.; Alzamil, I.; Almutairi, J. Enhanced Intrusion Detection in Drone Networks: A Cross-Layer Convolutional Attention Approach for Drone-to-Drone and Drone-to-Base Station Communications. *Drones* **2025**, *9*, 46. [[CrossRef](#)]
34. Zhang, Y.; Jin, B.; Zhu, Q.; Meng, Y.; Han, J. The effect of metadata on scientific literature tagging: A cross-field cross-model study. In Proceedings of the ACM Web Conference 2023, Austin, TX, USA, 30 April–4 May 2023; pp. 1626–1637.
35. Aldossary, M.; Almutairi, J.; Alzamil, I. Federated LeViT-ResUNet for Scalable and Privacy-Preserving Agricultural Monitoring Using Drone and Internet of Things Data. *Agronomy* **2025**, *15*, 928. [[CrossRef](#)]
36. Josse, J.; Chen, J.M.; Prost, N.; Varoquaux, G.; Scornet, E. On the consistency of supervised learning with missing values. *Stat. Pap.* **2024**, *65*, 5447–5479. [[CrossRef](#)]
37. Gharehchopogh, F.S. Quantum-inspired metaheuristic algorithms: Comprehensive survey and classification. *Artif. Intell. Rev.* **2023**, *56*, 5479–5543. [[CrossRef](#)]
38. Aldossary, M.; Alharbi, H.A.; Ayub, N. Exploring Multi-Task Learning for Forecasting Energy-Cost Resource Allocation in IoT-Cloud Systems. *Mathematics* **2024**, *79*, 4603–4620. [[CrossRef](#)]
39. Nayyef, H.M.; Ibrahim, A.A.; Mohd Zainuri, M.A.A.; Zulkifley, M.A.; Shareef, H. A novel hybrid algorithm based on jellyfish search and particle swarm optimization. *Mathematics* **2023**, *11*, 3210. [[CrossRef](#)]
40. Surya, S.; Muthukumaravel, A. Adaptive sailfish optimization-contrast limited adaptive histogram equalization (ASFO-CLAHE) for hyperparameter tuning in image enhancement. In *Computational Intelligence for Clinical Diagnosis*; Springer International Publishing: Cham, Switzerland, 2023; pp. 57–76.
41. Foody, G.M. Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient. *PLoS ONE* **2023**, *18*, e0291908. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.