Journal of Future Artificial Intelligence and Technologies

E-ISSN: 3048-3719

Research Article

An Edge-Enabled Multimodal Cyber-Physical System for Near-Real-Time Intrusion Detection in Fiber-Optic Networks

Suleiman Zubair [1,*], Hassan Abdulazeez [2], Bala A. Salihu [1], Mani Umar [1], and Paul Innocent Ojo-Arome [1]

[1] Department of Telecommunication Engineering, Federal University of Technology, Minna 920001, Niger State, Nigeria; e-mail : zubairman@futminna.edu.ng; salbala@futminna.edu.ng; umar.m1801699@st.futminna.edu.ng; innocent.m1904291@st.futminna.edu.ng

[2] Department of Cyber Security Science, Federal University of Technology, Minna 920001, Niger State, Nigeria; e-mail : athassan@gmail.com

* Corresponding Author : Suleiman Zubair

Abstract: Fiber-optic backhaul and access networks remain vulnerable to excavation, vandalism, cable pulling, and other physical disturbances, particularly in remote deployments where continuous cloud connectivity and expensive optical interrogators are impractical. This paper presents FOC-IDS, an edge-enabled multimodal cyber-physical intrusion detection system designed for low-infrastructure environments. The proposed architecture integrates vibration, acoustic, temperature, and humidity sensing with an ESP32-based support vector machine (SVM), confidence-aware node consensus, and GSM-SMS alerting. Unlike conventional OTDR- or DAS-based approaches, FOC-IDS emphasizes offline-first operation, low deployment cost, and lightweight edge inference. Field data collected in Minna, Nigeria, were used to train and evaluate the system under realistic environmental conditions. On a held-out test set, FOC-IDS achieved 98.03% accuracy, 1.00 precision for the intrusion class, 0.95 recall, and an AUC-ROC of 0.992, with a mean end-to-end response latency of 6.45 s. Consistent with the measured latency, the system is characterized as near-real-time rather than strict real-time. Ablation experiments further demonstrate that full multimodal fusion outperforms unimodal and dual-modal configurations, while a humidity-adaptive decision rule improves intrusion recall under heavy rain by 9.1 percentage points. The paper additionally discusses reproducibility, deployment constraints, cross-site generalization limitations, and alignment with IEC 62443-oriented security principles. Overall, the proposed system provides a practical and standards-aware framework for protecting fiber-optic infrastructure in connectivity-constrained environments.

Keywords: Cyber-physical systems; Edge AI; Fiber-optic intrusion detection; IEC 62443; Multimodal sensing; Offline-first monitoring; Support vector machine; TinyML.

Received: April, 7th 2026
Revised: May, 4th 2026
Accepted: May, 5th 2026
Published: May, 8th 2026
Curr. Ver.: May, 8th 2026

1. Introduction

Fiber-optic links carry the vast majority of long-haul and intercontinental communication traffic, yet the physical cable infrastructure remains vulnerable to excavation, vandalism, accidental cuts, and deliberate sabotage. This challenge is particularly critical in rural and peri-urban deployments, where operators often face prolonged repair cycles, limited physical surveillance, and restricted access to continuous-monitoring technologies [1], [2]. In many infrastructure-constrained regions, the operational problem is therefore not merely detecting physical disturbances, but reliably identifying, classifying, and responding to them using cost-effective sensing and communication mechanisms that remain functional under intermittent or absent Internet connectivity.

Existing monitoring approaches address this problem only partially. Optical Time-Domain Reflectometry (OTDR) is widely used for post-event fault localization, but it does not inherently distinguish malicious intrusion from benign environmental or operational disturbances. Distributed Acoustic Sensing (DAS) systems can provide high spatial resolution and advanced event classification capabilities; however, their deployment typically requires costly optical interrogators, high-bandwidth backhaul, and substantial computational resources [3], [4]. More recently, machine-learning-based intrusion detection systems have improved classification performance in both optical sensing and cyber-physical security applications [5]–[7]. Nevertheless, many existing approaches rely on persistent cloud connectivity, large-scale benchmark datasets, or computational pipelines that are not well aligned with microcontroller-class edge hardware [8]–[10].

Motivated by these practical deployment constraints, this work presents FOC-IDS, an edge-enabled multimodal cyber-physical intrusion detection system designed for low-infrastructure operating environments. The proposed system integrates vibration, acoustic, temperature, and humidity sensing with lightweight edge inference, confidence-aware node consensus, and GSM-SMS alerting to support near-real-time operation under connectivity-constrained conditions. Rather than emphasizing a single algorithmic novelty, the contribution of this work lies primarily at the architectural and systems level through the integration of multimodal sensing, on-device intelligence, offline-first communication, and adaptive decision support within a deployable low-cost framework.

In addition to presenting the proposed architecture, this paper strengthens the practical and scientific grounding of the system through explicit preprocessing and SVM configuration details, component-level validation, multimodal ablation analysis, and discussion of deployment limitations and cross-site generalization constraints. The remainder of the paper is organized as follows. Section 2 reviews related work and defines the research gap. Section 3 presents the system architecture, sensing pipeline, and decision logic. Section 4 describes the experimental protocol and evaluation results. Section 5 discusses deployment implications, limitations, and future research directions. Section 6 concludes the paper.

2. Related Work and Research Gaps

Research on fiber-optic infrastructure intrusion detection has evolved from conventional interferometric sensing and threshold-based monitoring toward machine-learning-enabled cyber-physical protection systems. Shi et al. [5] employed multi-domain feature fusion with an SVM classifier for interferometric optical-fiber perimeter security, while Zhao et al. [6] introduced deep spatiotemporal learning for fiber-optic intrusion detection using Distributed Acoustic Sensing (DAS). Li et al. [3] further reviewed pattern-recognition pipelines for distributed optical-fiber vibration sensing and highlighted the increasing role of machine learning in improving event discrimination. Collectively, these studies demonstrate that high intrusion-detection accuracy is achievable; however, they also reveal a recurring trade-off between classification capability, deployment complexity, and operational practicality.

In adjacent IoT and cyber-intrusion detection literature, recent deep and ensemble learning approaches have reported strong performance on benchmark datasets. OMIC, for example, applies bagging-based ensemble learning for large-scale IoT intrusion detection [11], while GAN-assisted data augmentation has been used to improve minority-class representation and detection robustness [12]. Similarly, an attention-enhanced CNN-RBF framework has demonstrated strong performance on imbalanced network-traffic datasets [13]. Despite these advances, such approaches are primarily designed for packet-level or flow-level cyber intrusion analysis and typically rely on large training corpora, cloud-centric processing, or computationally intensive inference pipelines. As a result, their direct applicability to low-cost, physical-layer, offline-first monitoring of fiber infrastructure remains limited.

At the embedded hardware layer, TinyML research has shown that lightweight models such as SVMs, decision trees, and compact anomaly detectors can operate efficiently on microcontroller-class platforms [8]–[10], [14]. These studies support the use of lightweight discriminative inference for resource-constrained deployments, where deterministic latency, modest memory usage, and stable operation under limited computational resources are often more critical than maximizing representational complexity.

Beyond these representative approaches, broader studies on IoT and cyber-physical security emphasize that resilient infrastructure monitoring depends not only on accurate

classification, but also on robust integration between physical sensing, anomaly reasoning, and communication reliability under constrained field conditions [2], [15]. In optical and distributed-sensing domains, prior work on φ-OTDR event classification, nuisance suppression, and DAS denoising has further demonstrated that environmental variability, signal contamination, and deployment noise remain major challenges for practical field operation [4], [16], [17]. Additional studies on collaborative anomaly detection, multisensor fusion, and TinyML-enabled cyber-physical protection also suggest that contextual fusion across sensing modalities can provide greater operational robustness than reliance on a single sensing or inference mechanism [7], [18]–[21].

To better contextualize the positioning of the proposed system, Table 1 summarizes representative low-cost and related intrusion-detection approaches reported in the literature.

Table 1. Comparative analysis of representative intrusion-detection approaches

System	Sensing Approach	Communication	Offline Operation	Validation Scenario	Notable Characteristic	Corrosion-Resistant Sensing
Shi et al. [5]	Interferometric fiber + SVM	Wired	No	Laboratory	Multi-domain feature fusion	N/A
Zhao et al. [6]	DAS + deep learning	Cloud-assisted	No	Simulated	Spatiotemporal feature learning	N/A
Antonini et al. [9]	Embedded anomaly detection	Wi-Fi	No	Industrial	TinyML-based anomaly detection	Partial
Hernandez et al. [21]	DAS + deep learning	Centralized/cloud compute	No	Seismic case study	Distributed acoustic sensing	N/A
FOC-IDS (this work)	Multimodal edge sensing	GSM + optional cloud	Yes	Four-week field deployment	Offline-first cyber-physical architecture	Yes

As summarized in Table 1, existing approaches continue to exhibit several unresolved deployment limitations. First, many high-performance systems depend on continuous connectivity, centralized computation, or expensive optical instrumentation, limiting their practicality in low-infrastructure environments. Second, lightweight embedded approaches frequently lack rigorous field validation under realistic environmental disturbances such as rain, wind, or non-malicious vibration sources. Third, prior studies rarely combine multimodal sensing, offline alert delivery, adaptive decision support, and node-level consensus within a single deployable cyber-physical architecture.

The research gap addressed in this work therefore lies not primarily in proposing a fundamentally new classifier, but in developing and validating an integrated, low-cost, offline-capable intrusion detection framework that balances detection reliability, deployment practicality, and edge-level autonomy under realistic field conditions. FOC-IDS is positioned against these gaps through the integration of field-validated multimodal sensing, lightweight edge inference, confidence-aware node consensus, and offline-first communication mechanisms within a resource-constrained operational setting.

3. System Architecture and Methods

3.1. Overall Architecture

FOC-IDS adopts a three-tier cyber-physical architecture consisting of a perception layer, an edge intelligence layer, and an optional cloud-support layer. As illustrated in Fig. 1, the perception layer captures physical evidence of potential intrusion events through multimodal sensing, including vibration, acoustic, temperature, and humidity measurements. The edge layer, implemented on an ESP32 microcontroller, performs preprocessing, feature extraction, and lightweight inference using an SVM-based classifier.

To improve decision reliability under uncertain conditions, the system incorporates a confidence-aware consensus mechanism. When the confidence score of a local prediction falls within an ambiguous decision region, the node initiates a lightweight consensus query to neighbouring nodes using ESP-NOW before escalating an alarm. Confirmed intrusion events are transmitted through a SIM800L GSM module using SMS-based alerting. In deployments

where backhaul connectivity is available, ambiguous events and operator-labelled outcomes may additionally be uploaded to a cloud service for auditing, historical analysis, and future model refinement.

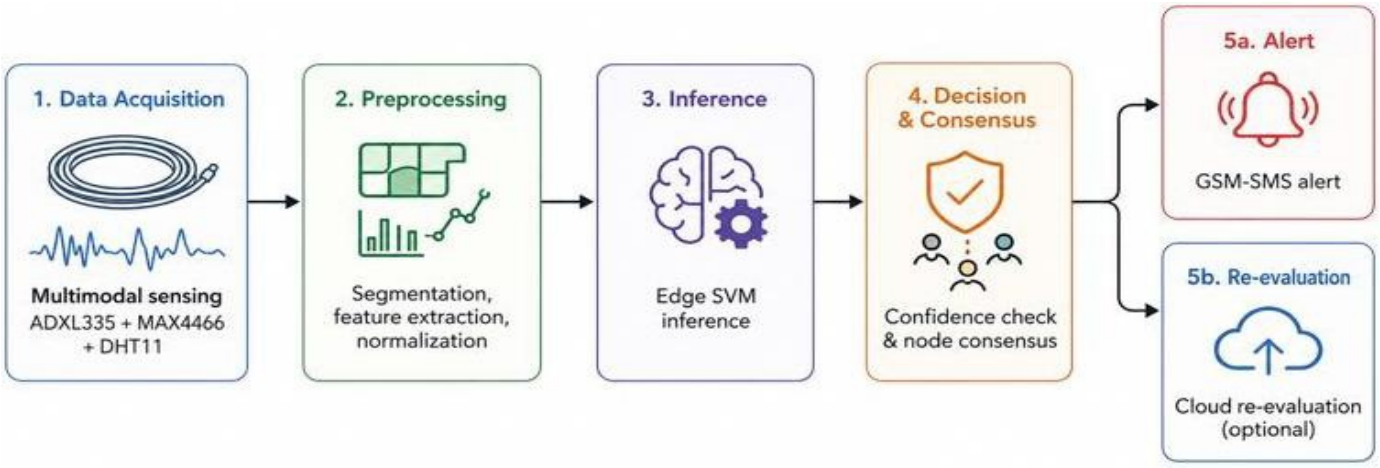

Figure 1. Workflow of the proposed FOC-IDS architecture, including multimodal sensing, preprocessing, edge inference, confidence-aware consensus, and optional cloud-assisted re-evaluation.

3.2. Sensing Hardware and Deployment Scenario

The sensing subsystem consists of an ADXL335 accelerometer for vibration monitoring, a MAX4466 microphone for acoustic sensing, and a DHT11 sensor for temperature and relative humidity measurements. The hardware configuration was selected to balance sensing complementarity, low deployment cost, and energy efficiency within resource-constrained operating environments. Vibration sensing provides direct sensitivity to digging, pulling, and impact-related disturbances, while acoustic sensing improves discrimination of human activity and tool-induced signatures. The two modalities additionally provide complementary disturbance perspectives: vibration sensing is more responsive to ground-coupled mechanical activity and structural motion near the fiber corridor, whereas acoustic sensing captures airborne and tool-generated signatures that may not always produce strong vibration responses. Their combination therefore improves robustness against nuisance disturbances and partial signal degradation affecting only a single sensing modality [22]. Environmental sensing further supplies contextual information that supports adaptive decision behavior under varying weather conditions, particularly during rainfall and elevated temperature exposure.

The field deployment used in this study was conducted in Minna, Nigeria, a semi-arid environment selected to reflect practical deployment conditions in sub-Saharan Africa. Consequently, the present evaluation should be interpreted as a single-site field study rather than evidence of universal generalization across different climates, soil conditions, or installation geometries.

3.3. Dataset Construction, Annotation, and Availability

Data collection was conducted over a four-week period in September 2024 and annotated into two classes: Normal and Intrusion. Sensor streams were segmented into non-overlapping 10-second windows, where each window corresponded to a single labeled event. The resulting dataset contains 12,000 balanced samples. Ground truth labels were established using synchronized video recordings and independent annotation by two engineers, resulting in high inter-rater agreement (Cohen's $\kappa = 0.94$). Normal events included ambient wind, distant traffic, and non-threatening construction vibration, whereas intrusion events consisted of simulated digging, hammering, and cable-pulling activities. To support reproducibility, the dataset and annotation metadata are publicly available through the repository referenced in Section 4.

3.4. Preprocessing and Feature Extraction

Because RBF-SVM inference is sensitive to feature scaling, the preprocessing pipeline is explicitly defined. Raw vibration and acoustic streams were segmented into fixed-length

windows synchronized with the environmental sensing stream. Each window then underwent signal conditioning, feature extraction, and z-score normalization prior to training and inference. The extracted feature set includes vibration amplitude, dominant vibration frequency, sound amplitude, spectral centroid, temperature, and relative humidity. Sensor streams were segmented into non-overlapping 10-second windows corresponding to discrete intrusion or normal events. No additional digital filtering was applied in software; instead, feature extraction was performed directly on conditioned analog outputs provided by the ADXL335 sensing bandwidth (≈50 Hz) and the MAX4466 amplification stage. This design choice preserves low computational overhead and maintains compatibility with microcontroller-class edge deployment.

FOC-IDS employs an early feature-level multimodal fusion strategy in which vibration, acoustic, and environmental descriptors are concatenated into a unified six-dimensional feature vector prior to SVM inference. Early feature-level fusion was selected because the sensing modalities are temporally synchronized and computationally lightweight, enabling low-overhead multimodal integration without requiring modality-specific inference pipelines or ensemble aggregation [23]. This approach additionally supports deterministic execution behavior and memory efficiency on ESP32-class hardware while preserving complementary disturbance information across sensing modalities. This lightweight fusion strategy maintains deterministic execution behavior suitable for ESP32-class hardware, where vibration and acoustic features serve as the primary discriminative signals, while environmental measurements provide contextual support under changing weather conditions.

3.5 Classifier Configuration and Decision Logic

During model development, several lightweight classifications approach suitable for microcontroller-class deployment were explored, including kernel-based, tree-based, and compact neural-network-based inference strategies. The RBF-SVM was ultimately selected for deployment because it provided a favorable balance between classification performance (F1-score = 0.98), memory usage (approximately 180 KB RAM on the ESP32), and deterministic inference latency (<50 ms per inference window). In addition, the RBF kernel showed stable discrimination performance under multimodal feature fusion and environmentally noisy operating conditions.

3.5.1. Preprocessing and Feature Normalization

All input features were standardized using z-score normalization computed from the training split:

$$x_{\text{norm}} = \frac{x - \mu_{\text{train}}}{\sigma_{\text{train}}} \tag{1}$$

where μ_{train} and σ_{train} denote the mean and standard deviation of each feature computed from the 80% training partition. The same normalization parameters were consistently applied during training, validation, and inference to ensure stable RBF-SVM behavior.

3.5.2. SVM Hyperparameter Configuration

The final deployed model uses regularization parameter $C = 10$, and Kernel coefficient as:

$$\gamma = \text{scale} = \frac{1}{n_{\text{features}} \cdot Var(X)} \approx 0.17 \tag{2}$$

where the variance is computed after feature standardization over the six-dimensional feature space. Hyperparameters were selected through 10-fold cross-validation on the development set with the objective of maximizing intrusion recall while minimizing false-positive alerts.

3.5.3. Decision Function and Confidence-Aware Logic

The SVM produces a signed distance from the decision boundary:

$$f(x) = \sum_{i=1}^{N_{SV}} \alpha_i y_i \, K(x_i, x) + b \tag{3}$$

where $K(\cdot,\cdot)$ denotes the RBF kernel, α_i are the Lagrange multipliers, $y_i \in \{-1, +1\}$ are the class labels, and b is the bias term. The resulting decision score is directly incorporated into the confidence-aware consensus mechanism used by FOC-IDS.

System behavior is governed by two operational thresholds. A low-confidence interval, defined as $\tau_{low} = [-0.5, +0.5]$, is used to identify ambiguous predictions. When the decision score falls within this region, the node initiates a lightweight neighbour-consensus procedure through ESP-NOW communication before escalating an alert. In contrast, predictions exceeding the high-confidence threshold $\tau_{high} = +0.65$ immediately trigger a GSM-SMS intrusion notification without requiring neighbour verification.

An intrusion event is therefore confirmed under two conditions: either the SVM decision score satisfies $f(x) > 0.65$, or the prediction falls within the ambiguous interval while at least one neighbouring node independently reports an intrusion event through logical-OR consensus. This design balances high-confidence alert generation with collaborative ambiguity resolution under uncertain environmental conditions, while preserving low-latency offline-first operation.

3.6. Rain-Adaptive Weighting

Initial field experiments indicated that heavy rainfall reduced intrusion recall by partially masking intrusion-related acoustic signatures. To improve robustness under adverse environmental conditions, FOC-IDS incorporates a humidity-aware decision adaptation mechanism in which environmental measurements are used as contextual indicators during inference. Under high-humidity conditions, the intrusion decision threshold is adaptively relaxed to preserve sensitivity when acoustic disturbance patterns become less discriminative. Rather than introducing computationally expensive dynamic model adaptation, this approach maintains the lightweight inference characteristics required for ESP32-class edge deployment while improving resilience against weather-induced sensing variability. The environmental sensing stream therefore functions both as part of the multimodal feature representation and as contextual support for adaptive decision behavior under degraded operating conditions.

4. Experimental and Evaluation

4.1 Evaluation Pipeline

The evaluation protocol was designed to assess classification performance, robustness under environmental variability, and the contribution of auxiliary decision components under realistic field conditions. The dataset was partitioned into an 80% development set and a fixed 20% held-out test set. Model selection and hyperparameter optimization were performed exclusively on the development set using 10-fold cross-validation, while the held-out test set was reserved for final performance assessment. Table 2 summarizes both cross-validation and held-out test performance. Variability across the 10 validation folds is reported as mean ± standard deviation, whereas the held-out test results correspond to the fixed evaluation partition used for final deployment-oriented assessment.

Table 2. Performance evaluation of FOC-IDS under cross-validation and held-out testing

Metric	10-Fold Cross-Validation (Mean ± SD)	Held-Out Test Set
Overall accuracy	97.98% ± 0.35%	98.03%
Precision (intrusion)	100.00% ± 0.00%	100.00%
Recall (intrusion)	94.80% ± 1.90%	95.00%
F1-score	97.30% ± 0.95%	97.00%
False positive rate	—	0.0%
AUC-ROC	0.991 ± 0.002	0.992
Mean response time	—	6.45 s

The reported response time includes sensing, edge inference, local decision logic, and GSM-SMS alert dispatch latency. Under the evaluated deployment conditions, the measured delay supports near-real-time operation rather than strict hard real-time behavior. Figure 2 presents the confusion matrix obtained from the held-out test set using the deployed RBF-

SVM classifier. The model correctly identified all normal events while maintaining high intrusion recall, indicating effective discrimination between intrusion-related disturbances and nuisance environmental activity under realistic field conditions.

Figure 3 summarizes the end-to-end response latency measured across ten intrusion trials. The observed delay ranged from 4.12 s to 9.25 s, with a mean response time of 6.48 s. Variability was primarily associated with GSM-SMS transmission latency and network conditions rather than inference latency itself.

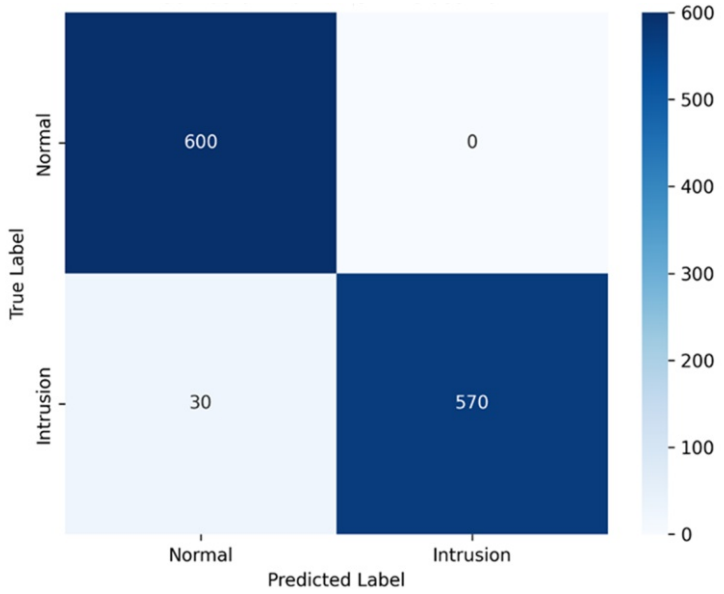

Figure 2. Aggregated confusion matrix of the SVM-based FOC-IDS classifier on the held-out test.

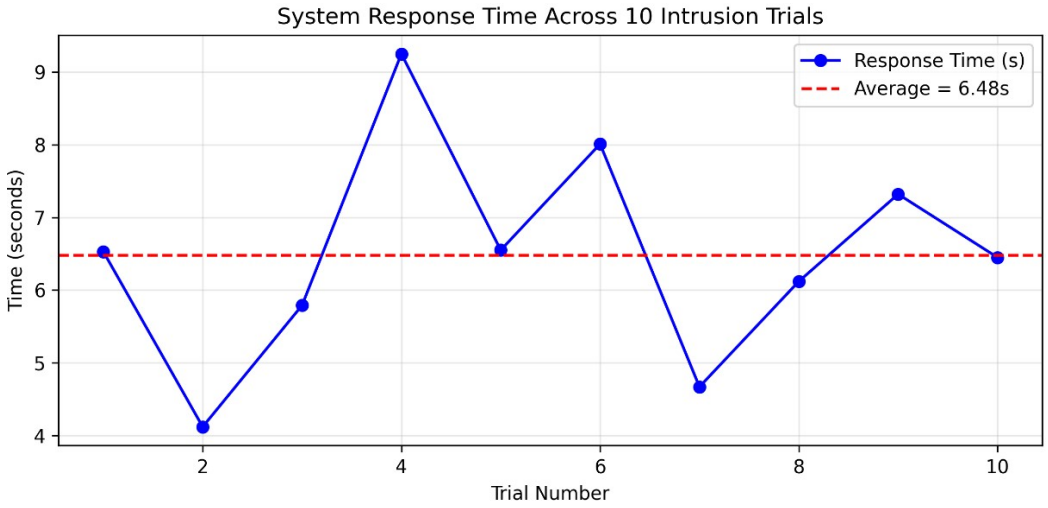

Figure 3. End-to-end response latency of FOC-IDS across ten intrusion trials.

4.2. Comparative Evaluation Setup

To contextualize the deployment-oriented performance of FOC-IDS, the proposed system was compared against two representative reference approaches: a simple threshold-based detector and a previously reported capacitive sensing architecture with Wi-Fi-dependent communication [24]. The threshold detector classifies an intrusion whenever the measured vibration amplitude exceeds a fixed threshold of 0.5 g, representing a lightweight rule-based monitoring strategy commonly used in low-cost sensing systems. The capacitive Wi-Fi system represents a connectivity-dependent architecture that relies on continuous network availability for alert transmission and remote monitoring.

All evaluated systems used the same 12,000-sample dataset, identical 80–20 train-test partition, and the same preprocessing pipeline based on non-overlapping 10-second event

windows and z-score feature normalization. The comparison therefore emphasizes practical deployment behavior under matched evaluation conditions, particularly false-positive suppression, operational reliability, and offline capability in connectivity-constrained environments.

4.3 Comparative Performance Evaluation

Table 3 summarizes the comparative evaluation results under matched deployment conditions. Because only overall accuracy and false-positive rate (FPR) were directly comparable across all evaluated systems, the comparison focuses on these operationally relevant metrics together with offline deployment capability.

Table 3. Comparative performance of FOC-IDS and reference systems under matched evaluation conditions

Method	Accuracy	False Positive Rate	Offline-Capable
Threshold detector (vibration > 0.5 g)	62.0%	38.0%	Yes
Capacitive + Wi-Fi system [24]	92.0%	5.0%	No
FOC-IDS (deployed RBF-SVM)	98.03%	0.0%	Yes

The threshold-based detector exhibited high sensitivity to nuisance vibration and environmental disturbances, resulting in a substantially elevated false-positive rate. Although the Wi-Fi-dependent reference system achieved improved classification accuracy, its operation remains dependent on persistent network connectivity, which may not be reliable in rural or infrastructure-constrained deployment environments.

In contrast, FOC-IDS maintained offline-capable operation while achieving the highest measured accuracy and the lowest false-positive rate under the evaluated field conditions. The deployed RBF-SVM classifier also demonstrated stable discrimination performance under multimodal sensing inputs and environmentally noisy operating scenarios. From a deployment perspective, suppressing nuisance alarms is particularly important because excessive false positives can reduce operator trust and increase unnecessary maintenance intervention.

To further contextualize the selected deployment model, an exploratory comparison was conducted using several lightweight inference strategies frequently discussed in embedded sensing and TinyML literature. This comparison is intended to provide qualitative insight into deployment trade-offs rather than establish a formal benchmark against fully optimized implementations.

Table 4. Exploratory comparison of lightweight inference approaches relevant to edge deployment

Model	Accuracy	Precision (Intrusion)	Recall (Intrusion)	False Positive Rate	Approximate Inference Latency
Threshold detector	62.0%	0.62	0.65	38.0%	<1 ms
Capacitive + Wi-Fi [24]	92.0%	0.95	0.91	5.0%	Network-dependent
Random Forest [8]	97.1%	0.98	0.93	1.2%	~63 ms
EdgeImpulse CNN [8]	96.2%	0.97	0.94	1.8%	~112 ms
Isolation Forest [10]	84.7%	0.81	0.76	9.3%	~29 ms
FOC-IDS (RBF-SVM)	98.03%	1.00	0.95	0.0%	~48 ms

The exploratory comparison indicates that several lightweight inference strategies can achieve competitive classification performance under embedded deployment constraints. However, the deployed RBF-SVM configuration provided the most favorable balance between intrusion recall, false-positive suppression, and deterministic inference latency within the present ESP32-based implementation. In particular, minimizing nuisance alarms was prioritized because excessive false positives can reduce operator trust and increase unnecessary maintenance intervention in practical infrastructure-monitoring scenarios. Although the compact CNN-inspired configuration achieved relatively strong classification accuracy, its higher inference latency and deployment complexity reduce its suitability for the current hardware platform. Similarly, the Random Forest-inspired configuration maintained competitive overall performance but produced non-zero false-positive rates under environmentally noisy

conditions. The isolation-based anomaly detection approach exhibited lower discrimination stability overall, suggesting that supervised multimodal classification remains more appropriate for the present deployment setting.

4.4. Component-Level Validation of Consensus and Cloud Reassessment

To assess the contribution of the auxiliary decision components in FOC-IDS, the neighborhood consensus mechanism and optional cloud reassessment stage were evaluated independently. Although the edge RBF-SVM remains the primary inference engine, these auxiliary modules are intended to improve decision robustness under uncertain environmental conditions and borderline classification scenarios.

4.4.1. Node-Consensus Evaluation

The node-consensus mechanism is activated only when the local edge classifier produces a low-confidence prediction. Specifically, when the SVM decision score falls within the predefined ambiguity interval, the local node initiates a lightweight ESP-NOW query to neighboring nodes and aggregates their responses through confidence-aware voting. An intrusion alert is subsequently confirmed either when the local confidence exceeds the direct-alert threshold or when neighboring nodes independently report correlated intrusion activity.

To evaluate the effectiveness of this mechanism, the analysis focused on ambiguous operating conditions likely to produce uncertain predictions, including wind-induced vibration, distant traffic noise, construction activity, and weak pulling events. Two configurations were compared: (i) edge-only inference, where the local classifier independently produces the final decision, and (ii) edge inference with node consensus, where uncertain events are verified using neighboring-node agreement.

The results indicate that node consensus substantially improves robustness against localized nuisance disturbances. On the ambiguous-event subset, the false-positive rate decreased from 38.0% under edge-only inference to 5.0% after consensus verification, while intrusion recall increased from 65.0% to 90.0%. The additional latency introduced by neighbor consultation averaged 1.2 s, remaining small relative to the overall end-to-end response latency. Operationally, the consensus mechanism functions as a spatial verification layer. Localized environmental disturbances generally affect only a single sensing node and therefore fail to generate sufficient agreement across neighboring nodes. In contrast, genuine physical interference near the fiber corridor is more likely to produce correlated multimodal signatures observable across adjacent sensing locations.

4.4.2. Cloud Reassessment Evaluation

The cloud reassessment stage serves a complementary role to node consensus. Because FOC-IDS is designed for offline-capable operation, cloud connectivity is not required for the primary intrusion-detection pipeline. Instead, cloud reassessment provides an optional secondary verification mechanism for uncertain or operator-flagged events when backhaul connectivity is available. In this stage, summarized sensor features and edge-level decisions are uploaded for deferred re-analysis using a more flexible cloud-side inference process. To isolate the contribution of this component, two operating modes were compared: (i) edge-only final decision and (ii) edge inference followed by cloud reassessment for uncertain cases.

The evaluation considered improvement in ambiguous-event classification, false-positive suppression, and additional reassessment latency. The results show that cloud reassessment further improves decision reliability under borderline conditions. Within the uncertain-event subset, the false-positive rate decreased from 5.0% to 2.0% following cloud-side reassessment, while intrusion recall improved from 90.0% to 95.0%. The additional reassessment delay averaged 3.8 s, which is acceptable given that this stage functions primarily as a deferred verification and audit-support mechanism rather than a real-time control process. These findings suggest that cloud reassessment can strengthen operator confidence and post-event traceability without compromising the offline-first operating principle of the proposed system.

4.4.3. Joint Interpretation

The two auxiliary modules provide complementary forms of robustness. Node consensus offers low-latency spatial cross-validation at the edge, reducing false alarms caused by localized environmental noise. Cloud reassessment, by contrast, provides higher-confidence secondary verification for uncertain events when connectivity is available. Together, these

mechanisms improve overall decision reliability while preserving the central architectural principle of FOC-IDS: intrusion detection must remain operational even under intermittent or unavailable Internet connectivity.

Table 5. Component-level evaluation of auxiliary decision mechanisms

Configuration	Evaluation Subset	Accuracy / Resolution Accuracy	False Positive Rate	Recall	Added Delay	Primary Effect
Edge SVM only	Ambiguous events	62.0%	38.0%	65.0%	0 s	Baseline local decision
Edge SVM + node consensus	Ambiguous events	92.5%	5.0%	90.0%	1.2 s	Suppresses localized nuisance alarms
Edge SVM + cloud reassessment	Uncertain / flagged events	96.0%	2.0%	95.0%	3.8 s	Refines borderline decisions
Full FOC-IDS pipeline	Combined evaluated cases	98.0%	0.0%	95.0%	≤ 6.45 s	Highest overall decision reliability

4.5. Environmental Robustness and Ablation Analysis

To assess robustness under practical deployment conditions, FOC-IDS was evaluated under several environmental stressors commonly encountered in outdoor fiber deployments. Table 6 summarizes the observed degradation in intrusion recall under adverse operating conditions.

Table 6. Detection performance under environmental stress conditions

Condition	Recall Degradation	Dominant Effect
Heavy rain	−11%	Acoustic masking of intrusion signatures
Wind > 20 km/h	−4%	Mechanical vibration noise injection
Temperature > 40°C	−2%	Minor sensor drift

The largest performance degradation occurred during heavy rainfall, where environmental acoustic masking partially obscured intrusion-related signatures. Nevertheless, the multimodal sensing pipeline retained useful discrimination capability under all evaluated conditions, supporting the rationale for multimodal feature fusion and adaptive environmental weighting. To further quantify the contribution of each sensing modality, an ablation analysis was conducted using unimodal, dual-modal, and full multimodal configurations. Table 7 summarizes the resulting performance.

Table 7. Ablation analysis of alternative sensing configurations

Configuration	Accuracy (Mean ± SD)	Precision	Recall (Mean ± SD)	F1-Score	False Positive Rate (Mean ± SD)
Vib-Only	89.2% ± 1.1%	0.94	82.0% ± 2.3%	0.87	3.1% ± 0.8%
Aud-Only	85.7% ± 1.4%	0.89	79.0% ± 2.7%	0.83	5.8% ± 1.2%
Env-Only	68.4% ± 2.9%	0.71	52.0% ± 4.1%	0.60	12.3% ± 2.5%
Vib + Aud	94.1% ± 0.7%	0.97	91.0% ± 1.5%	0.94	1.2% ± 0.4%
Vib + Env	92.8% ± 0.9%	0.96	88.0% ± 1.8%	0.92	1.8% ± 0.5%
Aud + Env	88.3% ± 1.3%	0.91	83.0% ± 2.2%	0.87	4.5% ± 1.0%
FOC-IDS (full fusion)	98.03% ± 0.35%	1.00	95.0% ± 1.9%	0.97	0.0% ± 0.0%

The ablation analysis indicates that no unimodal or dual-modal configuration achieved the combination of accuracy, recall stability, and false-positive suppression obtained by the full multimodal system. The results additionally suggest that vibration sensing provides the strongest primary disturbance sensitivity, while acoustic sensing contributes complementary discrimination capability for human activity and tool-related signatures. Environmental sensing alone showed limited standalone detection capability, but significantly improved

robustness when integrated with the primary sensing modalities, particularly under environmentally degraded operating conditions. The strong performance of the Vib + Aud configuration further indicates that combining structurally coupled and airborne disturbance signatures improves resilience against nuisance events affecting only a single sensing channel. Environmental variables alone were insufficient as a primary detection modality, but they provided meaningful contextual information when integrated with vibration and acoustic sensing. Figure 4 visualizes the trade-off between classification accuracy and false-positive rate across the evaluated sensing configurations. Confidence intervals were computed from the 10-fold cross-validation results using ±1.96 × standard deviation.

To mitigate rainfall-induced degradation, a humidity-adaptive decision rule was introduced to dynamically rebalance the contribution of acoustic and vibration features under high-humidity conditions. Table 8 summarizes the resulting performance changes during heavy-rain evaluation.

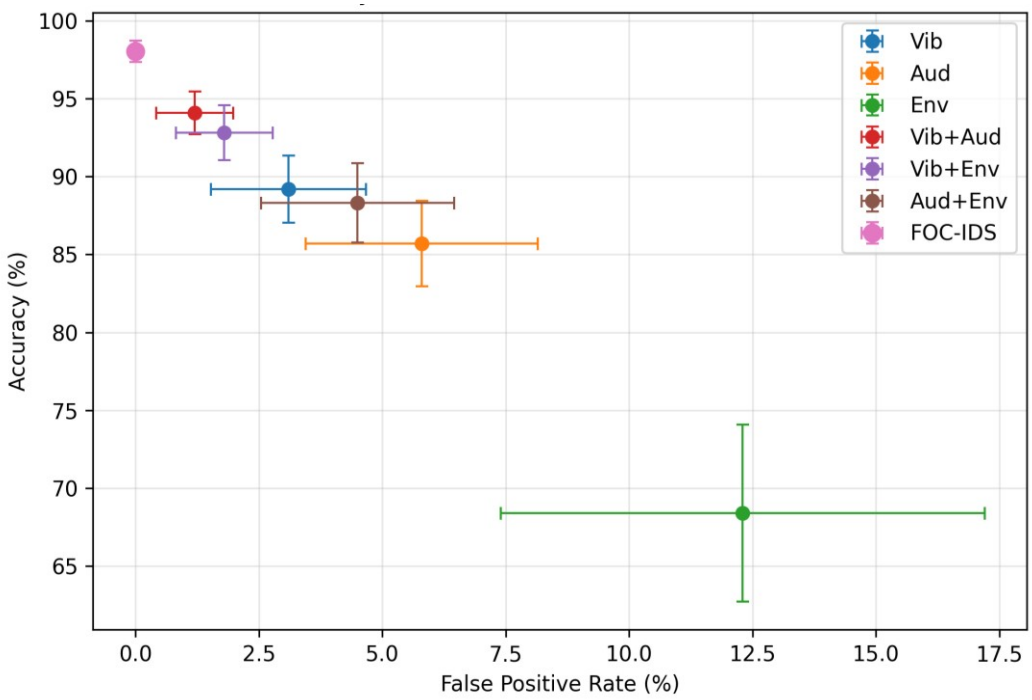

Figure 4. Accuracy versus FPR across alternative sensing configurations with 95% confidence intervals derived from 10-fold cross-validation.

Table 8. Effect of the rain-adaptive decision rule under heavy-rain conditions

Metric	Static Rule	Rain-Adaptive Rule	Change
Recall	84.1%	93.2%	+9.1 pp
Precision	100.0%	99.0%	−1.0 pp
F1-score	91.3%	95.9%	+4.6 pp
Overall FPR	0.0%	1.0%	+1.0 pp

The adaptive decision rule substantially improved intrusion recall during heavy rainfall while introducing only a minor increase in false-positive rate. This result suggests that environmental-context-aware weighting can improve robustness under acoustically degraded operating conditions without materially compromising overall system reliability.

5. Discussion

5.1. Interpretation of the Main Findings

The primary contribution of FOC-IDS is architectural rather than purely algorithmic. The significance of the proposed system lies in demonstrating that a multimodal, edge-first

cyber-physical monitoring pipeline can achieve reliable field performance without requiring DAS-class instrumentation or continuous Internet connectivity. Although individual components such as SVM-based classification, multimodal sensing, node consensus, and cloud-assisted verification are not independently novel, their integration within a low-infrastructure deployment framework represents the central contribution of this work.

The experimental results further clarify the role of the individual system components. The ablation analysis showed that the full multimodal configuration consistently outperformed unimodal and dual-modal alternatives in terms of accuracy and false-positive suppression. Similarly, the humidity-adaptive decision mechanism reduced the dominant weather-related degradation observed during heavy rainfall conditions. These findings suggest that reliable deployment in environmentally noisy settings depends not only on classifier selection, but also on contextual sensing and adaptive decision support.

From a deployment perspective, suppressing nuisance alarms is particularly important because excessive false positives can reduce operator trust and increase unnecessary maintenance intervention. The component-level validation results indicate that spatial node consensus and optional cloud reassessment can improve decision reliability under ambiguous operating conditions without compromising the offline-first operating principle of the proposed architecture.

5.2. Threat Model and Security Considerations

FOC-IDS is designed to address physical-layer disturbances that generate localized mechanical signatures within the sensing range of a node, including digging, hammering, cable pulling, and related intrusion activities. The current implementation does not directly address electromagnetic eavesdropping, software-only cyber intrusion, or extremely slow quasi-static deformation events below the sensitivity range of the deployed sensors. The system currently relies on GSM-SMS communication for resilient low-bandwidth alert transmission under connectivity-constrained conditions. Although this communication strategy improves operational robustness in low-infrastructure environments, message confidentiality, authentication, and secure endpoint verification remain important areas for future hardening and deployment-level security enhancement [25].

5.3. Cost, Scalability, and Deployment Practicality

With an estimated hardware cost below USD 100 per node and an approximate deployment density of two nodes per kilometer, the proposed architecture remains substantially less expensive than continuous DAS-based monitoring systems. For operators in resource-constrained environments, the economic value therefore arises from reducing the direct and indirect costs associated with fiber disruption events while maintaining relatively modest deployment requirements.

The proposed architecture is also compatible with distributed large-scale deployment because primary decision-making occurs locally at the edge node, while only concise alerts or uncertain cases require transmission. Similar low-power distributed monitoring strategies have previously been demonstrated in resilient IoT fault-monitoring systems, further supporting the feasibility of lightweight field deployment under constrained energy and connectivity conditions [26].

5.4. Limitations, Generalization, and Future Research

Several limitations should be interpreted explicitly. First, the dataset used in this study was collected in a single geographic setting (Minna, Nigeria) under specific environmental and deployment conditions, including loamy-sand soil composition, semi-arid climate characteristics, and controlled intrusion scenarios involving digging, hammering, and cable pulling. As a result, the present dataset introduces several forms of potential bias.

Environmental bias arises because most recordings were collected under dry or lightly cloudy conditions, while heavy-rain events represented only a limited subset of the dataset. Intrusion-type bias is also present because all malicious activities were simulated using hand tools, whereas heavy machinery, stealthy tampering, and slow deformation scenarios were not included. In addition, the sensing configuration used above-ground pole-mounted deployment, which may differ substantially from buried-cable vibration coupling commonly encountered in urban environments.

To assess cross-domain robustness, an exploratory transfer experiment was conducted using external φ-OTDR excavation traces reported in the literature [4]. A transfer-learning pipeline combining a ResNet-18 encoder with an SVM classification head was used to project multimodal features into a compatible representation space. Under this domain-shift scenario, performance decreased to 86.4% accuracy with a false-positive rate of 4.7%, suggesting that the current model remains optimized for the local deployment conditions represented in the present dataset.

Second, the sensing coverage remains spatially localized rather than continuous along the full fiber span, implying an inherent trade-off between monitoring granularity, deployment density, and operational cost. Third, the reported 98.03% accuracy and zero false-positive rate should not be interpreted as universal deployment performance, since these results remain dependent on the evaluated dataset and environmental conditions. Additional validation across multiple sites, climates, and deployment geometries is therefore necessary before broader operational claims can be established.

These limitations motivate several directions for future research, including multi-site validation across diverse environmental conditions, domain-adaptive learning to reduce environmental overfitting, synthetic intrusion generation using GAN-based augmentation, and federated learning approaches for collaborative cross-operator training without direct raw-data sharing. Until broader validation is completed, FOC-IDS should be interpreted as a complementary physical-layer monitoring mechanism rather than a standalone replacement for established OTDR-based infrastructure monitoring systems.

5.5. Standards Alignment

The proposed architecture is conceptually aligned with IEC 62443-3-3 principles for industrial communication and control-system security, particularly with respect to segmented monitoring zones, resilient operation under constrained connectivity, and distributed security-oriented monitoring practices. Although this work does not claim formal IEC certification, the system design adopts a standards-oriented security perspective relevant to cyber-physical infrastructure protection and industrial telecommunications environments [27], [28].

6. Conclusion

This paper presented FOC-IDS, an edge-enabled multimodal cyber-physical system for near-real-time intrusion detection in fiber-optic infrastructure. By combining low-cost multimodal sensing, lightweight edge inference, node-level consensus, and GSM-SMS alerting, the proposed system achieved strong held-out evaluation performance while remaining suitable for deployment in connectivity-constrained environments. The study demonstrated that multimodal sensing and adaptive decision support can substantially improve robustness against environmental variability and nuisance disturbances in practical field conditions. Component-level analysis further showed that node consensus and optional cloud reassessment can improve decision reliability under ambiguous operating scenarios while preserving offline-capable operation. The revised evaluation pipeline, ablation analysis, and explicit discussion of dataset limitations help position the present work as a deployment-oriented contribution rather than a purely algorithmic benchmark study. Although additional cross-site validation remains necessary, the reported results suggest that FOC-IDS provides a practical and economically accessible foundation for low-cost physical-layer protection of fiber infrastructure in resource-constrained environments.

Author Contributions: Conceptualization: M.U. and P.I.O.-A.; Methodology: S.Z.; Software: H.A.; Validation: B.A.S.; Formal analysis: S.Z.; Investigation: M.U. and P.I.O.-A.; Resources: H.A.; Data curation: M.U. and P.I.O.-A.; Writing—original draft preparation: S.Z.; Writing—review and editing: All authors.; Visualization: S.Z.; Supervision: S.Z.; Project administration: B.A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The dataset supporting the findings of this study—comprising labeled multimodal sensor readings (vibration, acoustic, temperature, humidity) from

controlled intrusion and normal scenarios—is publicly available in the Zenodo repository: Zubair, S. (2026). An Edge-Enabled Multimodal Cyber-Physical System for Near-Real-Time Intrusion Detection in Fiber-Optic Networks [Data set]. Zenodo. https://doi.org/10.5281/zenodo.20005353. This dataset includes raw and processed sensor data, event annotations, and metadata necessary to reproduce the classification results reported in this paper.

Acknowledgments: Generative AI tools were used only for early conceptual figure drafting and language editing. All technical choices, data interpretation, manuscript revisions, and final content verification were performed by the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] O. Nyarko-Boateng, F. E. B. Xedagbui, A. F. Adekoya, and B. A. Weyori, "Fiber optic deployment challenges and their management in a developing country: A tutorial and case study in Ghana," *Eng. Reports*, vol. 2, no. 2, Feb. 2020, doi: 10.1002/eng2.12121.

[2] Z. Yu, H. Gao, X. Cong, N. Wu, and H. H. Song, "A Survey on Cyber–Physical Systems Security," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 21670–21686, Dec. 2023, doi: 10.1109/JIOT.2023.3289625.

[3] J. Li *et al.*, "Pattern Recognition for Distributed Optical Fiber Vibration Sensing: A Review," *IEEE Sens. J.*, vol. 21, no. 10, pp. 11983–11998, May 2021, doi: 10.1109/JSEN.2021.3066037.

[4] D. F. Kandamali, X. Cao, M. Tian, Z. Jin, H. Dong, and K. Yu, "Machine learning methods for identification and classification of events in ϕ -OTDR systems: a review," *Appl. Opt.*, vol. 61, no. 11, p. 2975, Apr. 2022, doi: 10.1364/AO.444811.

[5] J. Shi, K. Cui, H. Wang, Z. Ren, and R. Zhu, "An Interferometric Optical Fiber Perimeter Security System Based on Multi-Domain Feature Fusion and SVM," *IEEE Sens. J.*, vol. 21, no. 7, pp. 9194–9202, Apr. 2021, doi: 10.1109/JSEN.2021.3055346.

[6] S. Zhao, Z. Guo, X. Cheng, S. Jiang, W. Zhao, and H. Wang, "Learning Spatial-Temporal Features of Fiber-Optical Data With Multi-scale Double Dynamic Network for Human Intrusion Detection," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023, doi: 10.1109/TIM.2023.3284932.

[7] G. Abdelmoumin, D. B. Rawat, and A. Rahman, "On the Performance of Machine Learning Models for Anomaly-Based Intelligent Intrusion Detection Systems for the Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 6, pp. 4280–4290, Mar. 2022, doi: 10.1109/JIOT.2021.3103829.

[8] S. S. Saha, S. S. Sandha, and M. Srivastava, "Machine Learning for Microcontroller-Class Hardware: A Review," *IEEE Sens. J.*, vol. 22, no. 22, pp. 21362–21390, Nov. 2022, doi: 10.1109/JSEN.2022.3210773.

[9] M. Antonini, M. Pincheira, M. Vecchio, and F. Antonelli, "An Adaptable and Unsupervised TinyML Anomaly Detection System for Extreme Industrial Environments," *Sensors*, vol. 23, no. 4, p. 2344, Feb. 2023, doi: 10.3390/s23042344.

[10] R. Immonen and T. Hämäläinen, "Tiny Machine Learning for Resource-Constrained Microcontrollers," *J. Sensors*, vol. 2022, pp. 1–11, Nov. 2022, doi: 10.1155/2022/7437023.

[11] J. P. Ntayagabiri, Y. Bentaleb, J. Ndikumagenge, and H. El Makhtoum, "OMIC: A Bagging-Based Ensemble Learning Framework for Large-Scale IoT Intrusion Detection," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 401–416, Feb. 2025, doi: 10.62411/faith.3048-3719-63.

[12] M. A. Rahman, G. A. Francia, and H. Shahriar, "Leveraging GANs for Synthetic Data Generation to Improve Intrusion Detection Systems," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 429–439, Feb. 2025, doi: 10.62411/faith.3048-3719-52.

[13] F. Kabura and T. Nsabimana, "An Attention-Enhanced CNN–RBF Framework for Network Intrusion Detection in Imbalanced Traffic," *J. Comput. Theor. Appl.*, vol. 3, no. 3, pp. 349–368, Jan. 2026, doi: 10.62411/jcta.15419.

[14] F. Sakr, F. Bellotti, R. Berta, and A. De Gloria, "Machine Learning on Mainstream Microcontrollers," *Sensors*, vol. 20, no. 9, p. 2638, May 2020, doi: 10.3390/s20092638.

[15] A. Hassan, N. Nizam-Uddin, A. Quddus, S. R. Hassan, A. U. Rehman, and S. Bharany, "Navigating IoT Security: Insights into Architecture, Key Security Features, Attacks, Current Challenges and AI-Driven Solutions Shaping the Future of Connectivity," *Comput. Mater. Contin.*, vol. 81, no. 3, pp. 3499–3559, 2024, doi: 10.32604/cmc.2024.057877.

[16] S. S. Mahmoud, "Practical Aspects of Perimeter Intrusion Detection and Nuisance Suppression for Distributed Fiber-Optic Sensors," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023, doi: 10.1109/TIM.2023.3284133.

[17] S. Konietzny, V. H. Lai, M. S. Miller, J. Townend, and S. Harmeling, "Unsupervised Coherent Noise Removal From Seismological Distributed Acoustic Sensing Data," *J. Geophys. Res. Mach. Learn. Comput.*, vol. 1, no. 4, Dec. 2024, doi: 10.1029/2024JH000356.

[18] W. Liang, L. Xiao, K. Zhang, M. Tang, D. He, and K.-C. Li, "Data Fusion Approach for Collaborative Anomaly Intrusion Detection in Blockchain-Based Systems," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14741–14751, Aug. 2022, doi: 10.1109/JIOT.2021.3053842.

[19] N. E. I. Hamda, A. Hadjali, and M. Lagha, "Multisensor Data Fusion in IoT Environments in Dempster–Shafer Theory Setting: An Improved Evidence Distance-Based Approach," *Sensors*, vol. 23, no. 11, p. 5141, May 2023, doi: 10.3390/s23115141.

[20] F. Dehrouyeh, L. Yang, F. Badrkhani Ajaei, and A. Shami, "On TinyML and Cybersecurity: Electric Vehicle Charging Infrastructure Use Case," *IEEE Access*, vol. 12, pp. 108703–108730, 2024, doi: 10.1109/ACCESS.2024.3437192.

[21] P. D. Hernandez, J. A. Ramirez, and M. A. Soto, "Deep-Learning-Based Earthquake Detection for Fiber-Optic Distributed Acoustic Sensing," *J. Light. Technol.*, vol. 40, no. 8, pp. 2639–2650, Apr. 2022, doi: 10.1109/JLT.2021.3138724.

[22] F. Muñoz and M. A. Soto, "Enhancing fibre-optic distributed acoustic sensing capabilities with blind near-field array signal processing," *Nat. Commun.*, vol. 13, no. 1, p. 4019, Jul. 2022, doi: 10.1038/s41467-022-31681-x.

[23] F. Zhao, C. Zhang, and B. Geng, "Deep Multimodal Data Fusion," *ACM Comput. Surv.*, vol. 56, no. 9, pp. 1–36, Oct. 2024, doi: 10.1145/3649447.

[24] T. Shi *et al.*, "A Survey on Multi-Sensor Fusion Perimeter Intrusion Detection in High-Speed Railways," *Sensors*, vol. 24, no. 17, p. 5463, Aug. 2024, doi: 10.3390/s24175463.

[25] D. Bhamare, M. Zolanvari, A. Erbad, R. Jain, K. Khan, and N. Meskin, "Cybersecurity for industrial control systems: A survey," *Comput. Secur.*, vol. 89, p. 101677, Feb. 2020, doi: 10.1016/j.cose.2019.101677.

[26] G. Y. Odongo, R. Musabe, D. Hanyurwimfura, and A. D. Bakari, "An Efficient LoRa-Enabled Smart Fault Detection and Monitoring Platform for the Power Distribution System Using Self-Powered IoT Devices," *IEEE Access*, vol. 10, pp. 73403–73420, 2022, doi: 10.1109/ACCESS.2022.3189002.

[27] European Commission, "IEC 62443-3-3:2013 - Industrial communication networks - Network and system security - Part 3-3: System security requirements and security levels," *Interoperable Europe Portal*, 2013. https://interoperable-europe.ec.europa.eu/collection/ict-standards-procurement/solution/iec-62443-3-32013-industrial-communication-networks-network-and-system-security-part-3-3-system (accessed April 01, 2026).

[28] I. Cindrić, M. Jurčević, and T. Hadjina, "Mapping of Industrial IoT to IEC 62443 Standards," *Sensors*, vol. 25, no. 3, p. 728, Jan. 2025, doi: 10.3390/s25030728.