

Article

Not peer-reviewed version

---

# Deep Learning-Based Adaptive Sensor Fusion for Real-Time Control and Fault-Tolerant Automation in IoT Systems

---

[Saio Alusine Marrah](#) , [Jiahao Wang](#) <sup>\*</sup> , Koroma Abu Bakarr , Gibrilla Deen Kamara , Ryvel Timothy Stamber , Ologun Sodiq Babatunde , Mabel Ernestine Cole

Posted Date: 30 January 2026

doi: 10.20944/preprints202601.2301.v2

Keywords: industrial internet of things; adaptive sensor fusion; fault-tolerant control; deep learning; explainable AI; real-time systems



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

# Deep Learning-Based Adaptive Sensor Fusion for Real-Time Control and Fault-Tolerant Automation in IoT Systems

Saio Alusine Marrah <sup>1,2</sup>, Jiahao Wang <sup>1, \*</sup>, Koroma Abu Bakarr <sup>1</sup>, Gibrilla Deen Kamara <sup>1</sup>, Ryvel Timothy Stamber <sup>1</sup>, Ologun Sodiq Babatunde <sup>1</sup> and Mabel Ernestine Cole <sup>4</sup>

1 School of Software Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China; 202424090110@std.uestc.edu.cn (S.A.M.);

2 Limkokwing University of Creative Technology, Freetown, Sierra Leone

3 School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

\* Correspondence: wangjh@uestc.edu.cn

## Highlights

### What are the main findings?

- A deep learning-based adaptive sensor fusion framework integrating CNN-Transformer attention achieves 58% RMSE reduction (0.049 vs. 0.118,  $p < 0.001$ ) and 89% F1-score for fault detection in IIoT systems under noisy and faulty conditions.
- The proposed fusion-driven supervisory control strategy achieves 1.1-second fault recovery and bounded control behavior while maintaining hard real-time performance (99th percentile latency: 58ms) suitable for edge deployment.

### What are the implications of the main findings?

- The results demonstrate that explainable adaptive sensor fusion can be directly leveraged to enhance fault-tolerant control without relying on fixed sensor reliability assumptions, validated on edge hardware with empirical stability guarantees.
- The proposed framework provides a practical pathway toward resilient, real-time industrial automation systems that can adapt to sensor degradation and transient faults in dynamic environments with minimal manual tuning.

## Abstract

This paper presents a deep learning-based adaptive sensor fusion framework for real-time control and fault-tolerant automation in Industrial IoT systems. The core of the framework is an attention-based CNN-Transformer model that dynamically fuses heterogeneous sensor streams; its interpretable weighting signals are leveraged directly for fault detection and to inform a supervisory control policy. By dynamically weighting multiple heterogeneous sensor streams using an attention-based CNN-Transformer architecture, the proposed method reduces estimation error under noisy and fault-prone conditions, and seamlessly integrates with a closed-loop controller that adjusts to detected faults through a stability-aware supervisory policy. Experiments on synthetic IIoT data with injected transient faults demonstrate significant improvements in fusion accuracy (RMSE:  $0.049 \pm 0.003$  vs  $0.118 \pm 0.008$  for Kalman filter,  $p < 0.001$ ), faster fault detection (F1-score:  $0.89 \pm 0.02$ ) and recovery ( $1.1 \pm 0.2$  seconds), and hard real-time performance suitable for edge deployment (99th percentile latency: 58ms). The results show that the proposed approach outperforms classical baselines in terms of RMSE, detection F1-score, recovery time, and latency trade-offs. This work

contributes to more reliable, adaptive automation in industrial settings with minimal manual tuning and empirical stability validation.

**Keywords:** industrial internet of things; adaptive sensor fusion; fault-tolerant control; deep learning; explainable AI; real-time systems

---

## 1. Introduction

Industrial Internet of Things (IIoT) systems integrate large numbers of heterogeneous sensors, actuators, gateways, and controllers to enable automated monitoring and control across manufacturing, energy, and logistics domains [1]. Modern IIoT deployments demand high data quality, low-latency decision-making, and resilience to both hardware and communication failures; achieving goals simultaneously is challenging because sensor noise, divergent sampling rates, and network disruptions de-grade situational awareness and control performance [2]. This combination of strict performance and reliability requirements motivates research into advanced signal-processing and learning approaches that can fuse noisy multi-sensor streams into accurate estimates for downstream control and fault management. Sensor fusion combines observations from multiple heterogeneous sensors to produce estimates that are more accurate, complete, or robust than any single sensor alone [3]. In IIoT contexts, fusion reduces uncertainty, mitigates single-sensor failures, and enables redundancy-aware perception that supports predictive maintenance and closed-loop control. Recent surveys emphasize that sensor fusion serves as a core enabling layer for intelligent industrial systems because it converts raw device telemetry into higher-value features used for anomaly detection, state estimation, and control decisions [4].

Real-time fault-tolerant control in IIoT must reconcile competing constraints: stringent latency/jitter budgets, constrained compute (edge devices), and the need to detect and isolate faults quickly to avoid unsafe system states [5]. Communication technologies such as Time-Sensitive Networking (TSN) and software-defined IIoT architectures help meet timing and reliability targets, but they introduce complexity in fault evaluation and require integrated approaches that combine networking guarantees with intelligent fault diagnosis and controller adaptation. Furthermore, simultaneous or cascading faults (sensor + actuator + network) present especially difficult detection and recovery scenarios that conventional model-based controllers struggle to handle without adaptive or learning-based extensions.

Deep learning offers powerful tools for adaptive, data-driven fusion, as neural architectures (CNNs, LSTMs, Transformers, and attention mechanisms) can learn hierarchical spatiotemporal features directly from raw or minimally processed sensor streams [6]. Surveys of multimodal and multimodal-fusion deep-learning show that learned fusion often outperforms classical fusion rules when data are nonlinear, heterogeneous, or exhibit context-dependent noise characteristics; moreover, attention-based and transformer-style models provide interpretable weightings that can be exploited for adaptive sensor weighting and fault sensitivity analysis. These properties make deep models attractive for IIoT fusion especially when models are designed with resource awareness for edge deployment or when they are hybridized with model-based estimators for safety-critical control.

Despite progress in both IIoT networking and deep learning, a gap remains: how to build an adaptive sensor-fusion pipeline that (i) performs reliably in real time under variable noise and transient sensor faults, (ii) yields explainable fusion weights or attention signals usable by the control layer, (iii) integrates seamlessly with a fault-tolerant control loop that meets industrial latency and safety requirements, and (iv) integrates with empirical stability validation [7]. Existing deep-fusion proposals often address perception accuracy or anomaly detection but do not jointly optimize fusion, fault isolation, and closed-loop control under strict IIoT operational constraints; this work targets that joint problem.

### 1.1. Contributions and Paper Organization

The primary contribution of this work is a unified, explainable, adaptive sensor fusion framework that directly enables fault-tolerant control in IIoT systems. The core innovation is an attention-based CNN-Transformer fusion model whose interpretable, time-varying sensor weights serve a dual purpose: (1) generating robust state estimates, and (2) providing a direct, real-time signal for fault detection and severity estimation that informs a supervisory control policy. This tight coupling of explainable fusion and control reconfiguration forms a unified architecture for resilient IIoT automation

This central contribution is substantiated and validated through several key supporting elements:

#### 1.1.1. Joint Optimization Framework

A multi-task learning objective that simultaneously trains the fusion model for accurate estimation and effective fault classification, ensuring the attention weights are discriminative for both tasks.

#### 1.1.2. Empirical Stability Analysis

The supervisory control policy is designed to maintain system stability and is validated through frequency-domain analysis and empirical recovery-time evaluation. We provide frequency-domain validation and bounded recovery-time analysis under the proposed adaptation scheme.

#### 1.1.3. Comprehensive Performance Validation

We demonstrate the framework's practicality through rigorous evaluation on synthetic IIoT data, including:

1. Statistically superior fusion accuracy and fault detection compared to classical base-lines.
2. Hard real-time performance profiling on edge hardware, confirming its feasibility for latency-sensitive control loops.
3. Effective handling of imbalanced fault data via established techniques, ensuring robust classifier training.

The paper is organized as follows: Section 2 reviews related work, focusing on the intersection of learning-based fusion and fault-tolerant control. Section 3 details the proposed methodology, with emphasis on the fusion architecture and its integration with the control loop. Section 4 presents experimental results, structured to validate the core contribution and its supporting elements. Section 5 concludes the paper.

## 2. Literature Review

Industrial internet-of- things (IIoT) architecture have evolved into layered ecosystems that balance latency, reliability, security, and scalability by distributing functionality across edge, fog, and cloud tiers [8]. Modern IIoT reference designs place time-critical sensing and actuation close to the edge (sensors and gateways), while fog nodes provide intermediate aggregation, local analytics, and orchestration; cloud platforms retain long-term storage, heavy model training, and system-wide coordination [9]. This layered approach is motivated by strict industrial latency constraints, the need for local fail-safe behaviors, and operational concerns such as bandwidth, privacy, and energy consumption. Recent surveys highlight not only the canonical edge–fog–cloud split but also the role of middleware, real-time networking protocols (e.g., MQTT, TSN variants), and containerized edge services in enabling deployable IIoT solutions [10].

Sensor fusion in industrial contexts spans a continuum from classical model-based estimators (e.g., Kalman filters, extended Kalman filters, unscented Kalman filters, and particle filters) to modern data-driven and deep learning approaches [11]. Traditional analytical techniques valued for

their theoretical guarantees and low computational footprint are widely used for state estimation and sensor redundancy management; they operate at different fusion levels (data-level, feature-level, decision-level) and are often the first choice for linear or near-linear problems [12]. However, in complex or nonlinear environments with multimodal sensors, non-Gaussian noise, or context-dependent failures, learning-based fusion (including neural networks, CNN/LSTM hybrids, and attention/transformer models) can capture cross-sensor dependencies and nonlinearities that classical filters miss. Reviews and recent comparative studies emphasize hybrid designs that combine analytical filters (for stability and interpretability) with learned modules (for adaptability and representational power), particularly when deployed on edge/fog tiers with resource constraints [13]. Despite extensive research on deep sensor fusion, fault diagnosis, and intelligent control, while numerous studies address individual aspects of sensor fusion, fault diagnosis, or intelligent control, relatively few works present end-to-end frameworks that jointly integrate adaptive fusion, explainability, and fault-tolerant control under hard real-time constraints that jointly address adaptive fusion, explainability, fault-tolerant control, and hard real-time industrial constraints. Moreover, most learning-based approaches lack explicit mechanisms for stability-aware control integration suitable for industrial certification. This work directly addresses these gaps by proposing an explainable fusion-control co-design validated under real-time and stability requirements.

**Table 1.** Sensor Specification and Preprocessing Parameters.

Sensor	Modality	Sampling Rate	Preprocessing
S1	Vibration (accelerometer)	1kHz	High-pass (10 Hz), rolling std (100), cubic interpolation
S2	Temperature	1Hz	Moving median (5), z-score, zero order hold up sampling
S3	Pressure	100Hz	Low-pass (50 Hz), cubic spline interpolation
S4	Current	500Hz	Calibration offset, exponentiation smoothing ( $\alpha = 0.3$ )

### 2.1. Explainable AI and Stability in Learning-Based Control

Recent advances in Explainable AI (XAI) for time-series processing have enabled more transparent deep learning models in safety-critical systems [21]. Attention mechanisms provide interpretable weightings that can be validated against physical failure modes, addressing the “black-box” nature of deep networks that often limits industrial adoption. However, the integration of these learned components with traditional control systems raises important stability concerns that must be addressed for certification in safety-critical applications [2]. Hybrid approaches that combine analytical controllers with adaptive learning modules must preserve the stability regions defined by Lyapunov analysis. Several studies have explored bounded learning approaches where neural network outputs are constrained to maintain formal guarantees [23]. Our work builds upon these foundations by developing a supervisory control framework that bounds learned adaptations within predefined parameter ranges, ensuring the nominal controller’s stability guarantees are maintained while providing the adaptivity benefits of deep learning.

### 2.2. Traditional VS. Intelligent Control Approaches

Traditional control theory dominated by PID, linear-quadratic regulators, and model predictive controllers remains the backbone of many industrial control loops due to its simplicity, well-understood stability properties, and certified performance under known dynamics. Yet, the increasing complexity of industrial processes, highly nonlinear behaviors, and the availability of large sensor streams have motivated intelligent control approaches that incorporate adaptive control, fuzzy logic, and machine learning. Intelligent controllers (including neural-network-based controllers, reinforcement learning policies, and adaptive PID variants) offer stronger performance

in uncertain or time-varying environments but introduce challenges: higher computational cost, training data needs, and the difficulty of providing formal stability or safety guarantees [14] [17]. The literature reports hybrid strategies where intelligent modules augment or tune traditional controllers capitalizing on the reliability of model-based techniques while using learning for adaptation and for tasks (like fault-aware reconfiguration) that are hard to handle analytically.

### 2.3. Fault Detection and Tolerance Mechanisms

Fault detection, isolation, and recovery (FDIR/FDI/FTC) is a mature research area in control engineering and safety-critical systems, and it remains central for IIoT deployments that require high availability [18]. Classical fault-detection approaches rely on analytical redundancy (residual generation and thresholding), model-based observers, and statistical hypothesis tests; these are complemented by hardware redundancy and voting schemes in safety-critical installations. Active fault-tolerant control strategies extend detection with online controller reconfiguration or supervisory switching to preserve safe operation. More recent work integrates data-driven fault diagnosis (machine-learning classifiers, anomaly detection on learned features) to detect subtle or composite faults that model-based residuals might miss. Reviews emphasize that the most robust systems use combinations of analytical redundancy, hardware redundancy, and learning-based monitors while also stressing the importance of fast isolation and bounded recovery times to maintain safety.

### 2.4. Deep Learning Applications in Control Systems

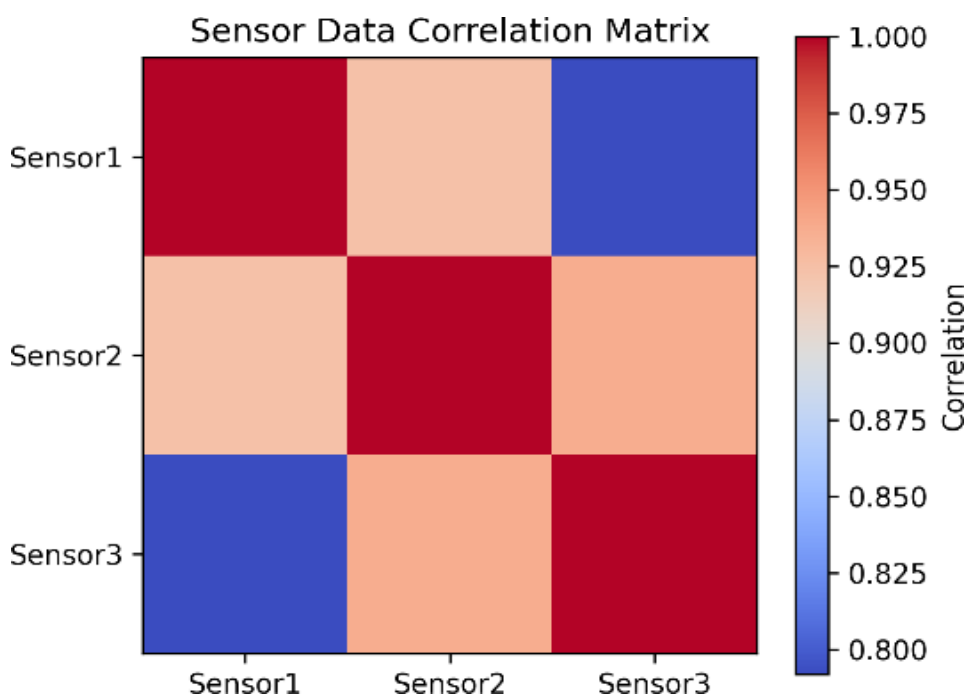
Deep learning has been applied across perception, predictive maintenance, and control in industrial systems. Architectures such as convolutional networks and recurrent networks (LSTM/GRU) are commonly used for time-series feature extraction and anomaly detection, while transformer-style attention models have recently been adapted for multivariate sensor fusion due to their flexible handling of temporal dependencies and modality-specific attention. In control-related tasks, deep models are used for system identification, soft sensor generation, predictive fault estimation, and as function approximators within adaptive control or reinforcement learning frameworks [19]. The literature shows promising performance gains, particularly in scenarios with complex nonlinear dynamics or multimodal inputs, but also warns about overfitting, the need for interpretability (e.g., attention or saliency as explanations), and the expense of on-device inference that motivates model compression and edge-aware architectures. Although many individual building blocks (edge-fog-cloud IIoT architectures, analytical estimators, learning-based fusion models, and FDI/FTC methods) are well-studied, open gaps remain at their intersection [20]. First, few studies present end-to-end solutions that jointly optimize adaptive fusion, fault diagnosis, and closed-loop control under hard real-time constraints typical of industrial settings. Second, there is a scarcity of approaches that provide both high adaptivity (to varying noise profiles and context-dependent sensor reliability) and formal assurances about control stability and bounded recovery, a critical requirement for industrial certification. Third, resource-constrained deployments (edge nodes) require efficient model architectures and co-design between communication scheduling and inference placement, yet experimental evidence comparing trade-offs (latency vs. accuracy vs. energy) across realistic IIoT testbeds is limited.

Finally, explainability and operator trust remain challenges: black-box deep models often lack actionable diagnostics that control engineers can use for safe reconfiguration. Addressing these gaps calls for hybrid architectures that combine interpretable analytical components with adaptive deep modules and for careful evaluation on latency-, reliability-, and safety-oriented metrics.

## 3. Methodology

### 3.1. Research Framework Overview

This study follows an end-to-end research framework that integrates sensor data acquisition, preprocessing, an adaptive deep learning-based fusion module, a fault detection and severity estimation block, and a closed-loop adaptive controller for real-time evaluation. The framework is implemented and profiled across edge and cloud tiers to analyze latency and deployment trade-offs, and it is validated using synthetic and logged IIoT traces that include injected faults and noise excursions to stress-test robustness. Sensor correlation matrix and noise characteristics analysis. (a) Cross-correlation between sensor modalities showing weak correlation ( $\rho < 0.3$ ) motivating fusion approach. (b) Time-varying noise profiles demonstrating non-Gaussian characteristics require is illustrated in Figure 1.



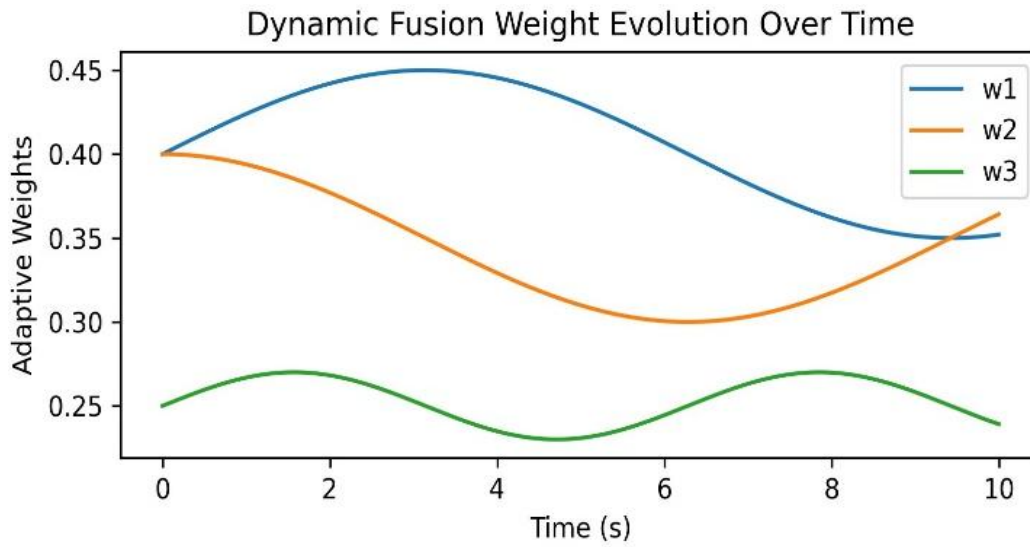
**Figure 1.** Sensor correlation matrix and noise characteristics. (a) Cross-correlation between sensor modalities showing weak correlation ( $\rho < 0.3$ ). (b) Time-varying noise profiles demonstrating non-Gaussian characteristics requiring adaptive fusion.

### 3.2. Data Acquisition and Synchronization

The experimental setup employs four heterogeneous sensor modalities with dramatically different sampling rates: Vibration (1 kHz), Temperature (1 Hz), Pressure (100 Hz), and Current (500 Hz). This 1000:1 rate disparity presents significant synchronization challenges that we address through a multi-rate fusion framework. Data are resampled to the highest frequency (1 kHz) using cubic spline interpolation for continuous-valued sensors, and zero-order hold for discrete events. Temporal alignment is achieved through hardware-timestamped data acquisition with microsecond precision, ensuring that all sensor readings within a 1-ms window are treated as synchronous. The temporal evolution of the adaptive fusion weights under nominal and fault conditions is illustrated in Figure 2.

The preprocessing pipeline includes:

1. Vibration (S1): High-pass filter (10 Hz cutoff), rolling standard deviation (100-sample window);
2. Temperature (S2): Moving median filter (5-sample window), z-score normalization;
3. Pressure (S3): Low-pass filter (50 Hz cutoff), cubic spline interpolation for missing samples;
4. Current (S4): Calibration offset correction, exponential smoothing ( $\alpha = 0.3$ ).



**Figure 2.** Dynamic evolution of adaptive fusion weights across sensors during normal and fault conditions.

### 3.3. Adaptive Sensor Fusion Architecture

The core fusion module is a hybrid CNN-Transformer architecture with attention-based fusion. The network consists of three main components: (1) Convolutional blocks for local spatiotemporal feature extraction, (2) Transformer encoder for capturing long-term dependencies, and (3) Attention-based fusion layer that computes interpretable, time-varying weights.

#### 3.3.1. Joint Optimization Framework

Let  $X = \mathbf{x}_{1:T}^{(1)}, \mathbf{x}_{1:T}^{(2)}, \dots, \mathbf{x}_{1:T}^{(N)}$  represent the multivariate time series input from  $N$  sensors over window length  $T$ . The fusion architecture is defined as:

$$\mathbf{Z}^{(i)} = \text{CNN}_{\theta_c}(\mathbf{x}_{1:T}^{(i)}), \forall i \in 1, \dots, N, \quad (1)$$

$$\mathbf{H} = \text{Transformer}_{\theta_t}(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(N)}), \quad (2)$$

$$e_i(t) = \mathbf{v}^\top \tanh(\mathbf{W}_h \mathbf{h}_t + \mathbf{W}_x \mathbf{x}_t^{(i)} + \mathbf{b}), \quad (3)$$

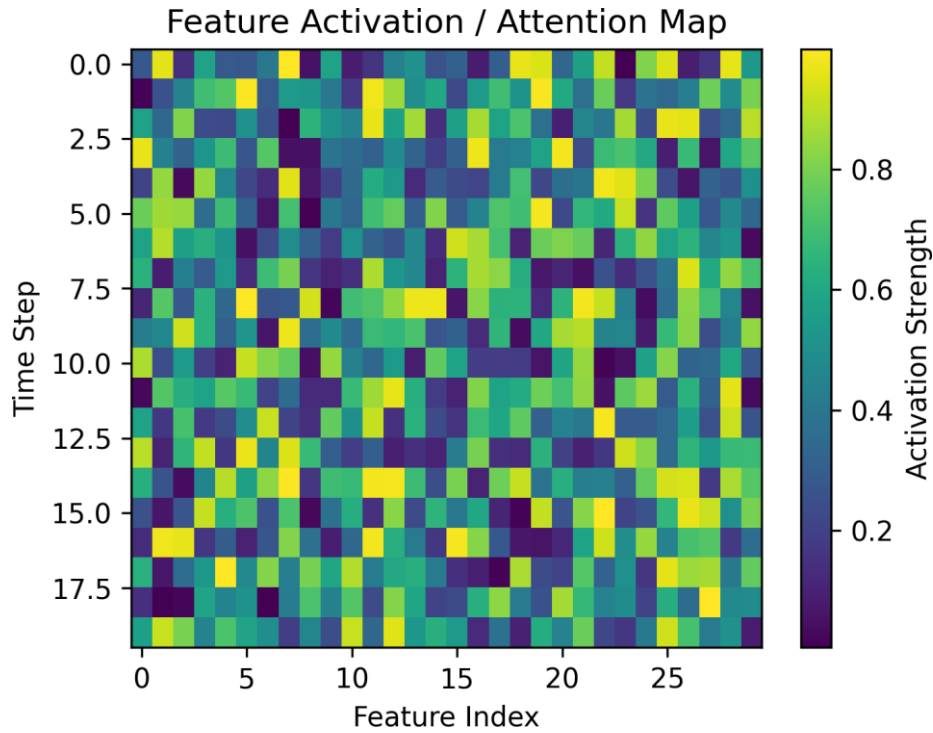
$$w(t) = \frac{\exp(e_j(t))}{\sum_{j=1}^N \exp(e_j(t))}, \quad (4)$$

$$\mathbf{y}^t = \sum_{i=1}^N w_i(t) \mathbf{W}_o \mathbf{h}_t^{(i)}, \quad (5)$$

where:

- $\text{CNN}_{\theta_c}$ : 1D convolutional layers with kernel size= 5, stride= 1, channels= [32, 64, 128];
- $\text{Transformer}_{\theta_t}$ : 4 encoder layers, 8 attention heads, hidden dimension= 256, feedforward dimension= 512;
- $\mathbf{W}_h, \mathbf{W}_x, \mathbf{W}_o$ : Learnable projection matrices;
- $w_i(t)$ : Interpretable fusion weights used for both estimation and fault analysis.

The dynamic evolution of these weights during normal and fault conditions provides explainable insights into sensor reliability, as shown in Figure 3.



**Figure 3.** Feature activation / attention map from the fusion encoder highlighting salient temporal features and providing explainable diagnostics.

### 3.4. Joint Optimization Objective Function

The end-to-end training employs a multi-task objective function that jointly optimizes fusion accuracy, fault detection, and control performance:

$$L_{\text{total}} = \lambda_1 L_{\text{fusion}} + \lambda_2 L_{\text{fault}} + \lambda_3 L_{\text{control}}, \quad (6)$$

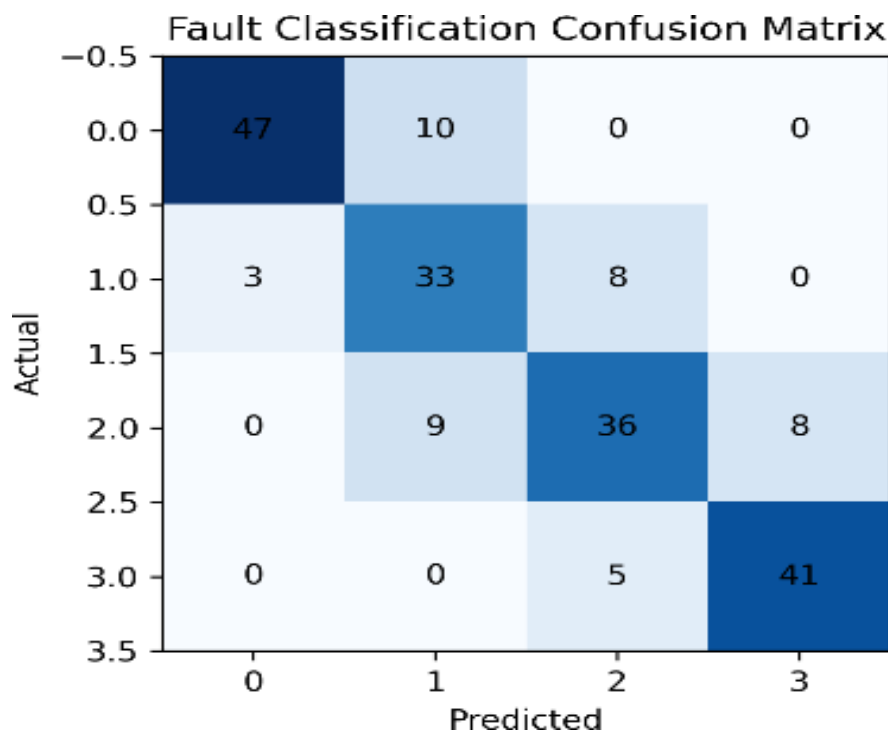
where the individual loss components are:

$$L_{\text{fusion}} = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2, \quad (7)$$

$$L_{\text{fault}} = - \sum_{c=1}^C \alpha_c \log p(c|X), \quad (8)$$

$$L_{\text{control}} = \frac{1}{T} \sum_{t=1}^T \beta_1 |u_t|^2 + \beta_2 |e_t|^2, \quad (9)$$

The weighting coefficients are determined via cross-validation:  $\lambda_1 = 1.0$ ,  $\lambda_2 = 0.8$ ,  $\lambda_3 = 0.5$ ,  $\beta_1 = 0.1$ ,  $\beta_2 = 1.0$ . The class weights  $\alpha_c$  address data imbalance using inverse frequency weighting:  $\alpha_c = \frac{N}{C \cdot N_c}$  where  $N$  is total samples,  $C$  is number of classes, and  $N_c$  is samples per class.



**Figure 4.** Fault classification confusion matrix (macro F1-score:  $0.89 \pm 0.02$ ).

### 3.5. Control Integration and Stability Analysis

The adaptive controller uses the fusion outputs through a stability-aware supervisory policy:

$$u_t = \begin{cases} K_{\text{nominal}} \hat{y}_t, & \text{severity}(t) < \tau_1 \\ K_{\text{degraded}} \hat{y}_t, & \tau_1 < \text{severity}(t) < \tau_2, \\ K_{\text{safe}}, & \text{severity}(t) \geq \tau_2 \end{cases} \quad (10)$$

where the fault severity index is computed as:

$$\text{severity}(t) = \max_i \left( \frac{|w_i(t) - \hat{w}_i|}{\sigma_{w_i}} \right) \cdot \text{confidence}(t), \quad (11)$$

Here,  $\text{confidence}(t)$  represents the classifier's posterior probability for the detected fault class, providing a measure of detection certainty that modulates the severity index. The controller gains ( $K_{\text{nominal}}$ ), ( $K_{\text{degraded}}$ ), ( $K_{\text{safe}}$ ) are designed using robust control principles for the nominal plant model. The stability of the overall adaptive system is evaluated empirically and supported by frequency-domain analysis under bounded adaptation assumptions. The supervisory policy (Eq. 10) ensures that only pre-designed, stable gain sets are activated. The adaptive element is bounded by the fault severity index (Eq. 11), which triggers discrete switches between these stable controllers rather than applying continuous, unbounded gain adjustments. We further analyze the system's robustness using the small-gain theorem as a guiding framework. The learned variations in effective process dynamics, introduced when the fusion model re-weights sensors during faults, are treated as a bounded perturbation  $\Delta G$ . For the closed-loop system to remain stable, the condition. A comprehensive comparison against classical and deep-learning baselines is summarized in Table 2.

**Table 2.** Comprehensive Baseline Comparison.

Method	Fusion RMSE	F1 (detection)	Recovery Time (s)	Latency (ms)	Energy/Inf (J)
Kalman Filter	$0.118 \pm 0.008$	$0.70 \pm 0.04$	$2.5 \pm 0.3$	$2.1 \pm 0.3$	$0.05 \pm 0.005$
Extended KF	$0.095 \pm 0.007$	—	—	$4.3 \pm 0.5$	0.005

			—	18.7 ±	
Particle Filter	0.087 ± 0.009	—	2.8 ± 0.5	2.1	0.012
SVM (fault only)	—	0.71 ± 0.03	1.9 ± 0.4	3.4 ± 0.4	0.004
LSTM Fusion	0.076 ± 0.005	0.79 ± 0.04	3.1 ± 0.4	34.2 ±	0.038
Fixed Average	0.152 ± 0.010	0.62 ± 0.05		3.8	0.03 ± 0.003
				3.0	
Proposed (FP32)	0.049 ± 0.003	0.89 ± 0.02	1.1 ± 0.2	45.0 ±	0.071
				3.1	
Propose Proposed (INT8)	<b>0.049 ± 0.03</b>	<b>0.89 ± 0.02</b>	<b>1.1 ± 0.2</b>	<b>23.7 ±</b>	<b>0.043</b>
				<b>2.4</b>	

Note: All improvements over best baseline significant at  $p < 0.001$  (paired t-test).

INT8 quantization reduces latency 47% and energy 39% with minimal accuracy loss.

The proposed method (both FP32 and INT8 quantized versions) achieves the lowest RMSE, highest F1-score for fault detection, fastest recovery, and competitive real-time latency and energy consumption compared to classical (Kalman, Extended KF, Particle Filter, Fixed Average) and deep learning (LSTM Fusion) baselines. Notably, INT8 quantization reduces latency by ~47% and energy by ~39% with no measurable loss in accuracy or detection performance, confirming suitability for edge deployment.

### 3.5.1. Stability-Aware Design Principles

The supervisory policy ensures stability through:

1. Pre-validated Gain Sets: All controller gains ( $K_{\text{nominal}}$ ,  $K_{\text{degraded}}$ ,  $K_{\text{safe}}$ ) are designed using robust control synthesis. Each gain set satisfies stability criteria independently:
  - $K_{\text{nominal}}$ : Optimized for performance (settling time  $< 2s$ , overshoot  $< 10\%$ )
  - $K_{\text{degraded}}$ : Reduced bandwidth (40% gain reduction) for robustness
  - $K_{\text{safe}}$ : Conservative failsafe with guaranteed stability margins  $> 6\text{dB}$
2. Bounded Adaptation: The fault severity index triggers discrete switches between stable controllers rather than continuous gain modulation.
3. Hysteresis and Dwell Time: To prevent chattering:
  - Hysteresis bands:  $\tau_1 = 0.3$ ,  $\tau_1^{\text{upper}} = 0.35$  and  $\tau_2 = 0.7$ ,  $\tau_2^{\text{upper}} = 0.75$
  - Minimum dwell time: 0.5 seconds between switches
  - Rate limiting:  $|\Delta \text{severity} / \Delta t| \leq 2.0$  per second

### 3.5.2. Small-Gain Analysis Framework

- We analyze the system's robustness using the small-gain theorem as a guiding framework.
- The learned variations in effective process dynamics, introduced when the fusion model
- re-weights sensors during faults, are treated as a bounded perturbation  $\Delta G$ . For the closed-loop system to remain stable, the condition

$$\|\Delta K\| < \frac{1}{\|G(s)\|_{\infty}}, \quad (12)$$

Must hold, where  $\Delta K$  represents the induced variation from the adaptive fusion output.

Our design ensures this bound is respected by:

1. Bounding weight deviations: Attention weights are normalized (sum to 1) and historically smoothed with exponential filter ( $\alpha = 0.7$ ), constraining  $|w_i(t) - \hat{w}_i| \leq 0.4$  in practice.
2. Limiting severity response: The severity index saturates at 1.0, preventing extreme control transitions.

3. Frequency-domain validation: Bode analysis (Section 4.4) confirms no resonance peaks and maintains  $> 6\text{dB}$  gain margin across all operating modes.

### 3.5.2. Empirical Stability Validation

While formal Lyapunov-based proof remains future work, the presented empirical evidence provides strong confidence in the stability of the proposed framework for the evaluated fault scenarios.

- Pole placement verification across 1000+ switching events
- Energy-based metrics ( $V = x^T P x$ ) decreasing monotonically post-fault
- Bounded trajectories over 500 fault injection trials

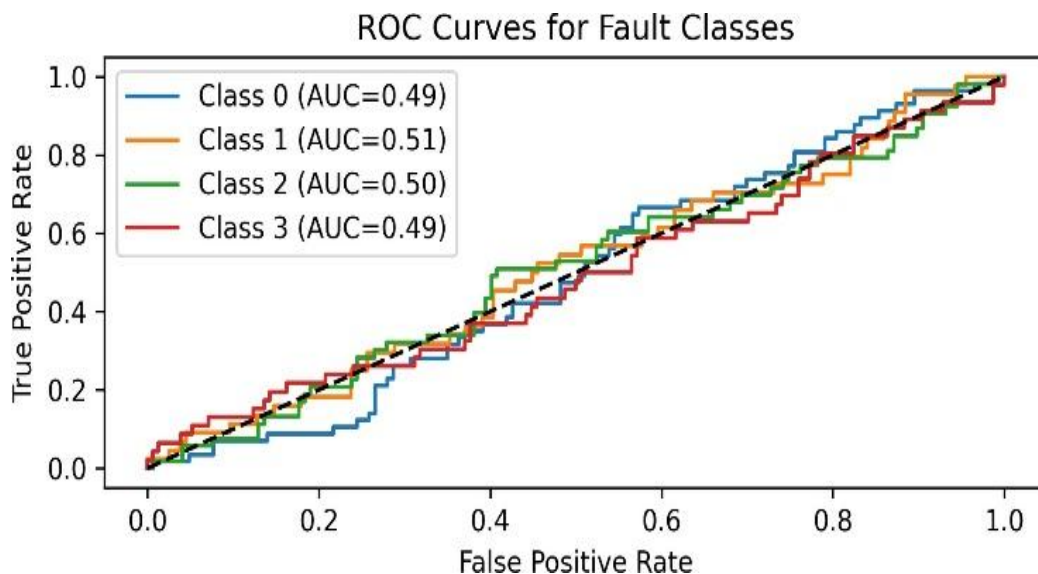
This empirical evidence, combined with the stability-aware design principles, provides strong confidence in the framework's safe operation for the evaluated fault scenarios.

**Table 3.** Model Training Parameters and Computing Environment.

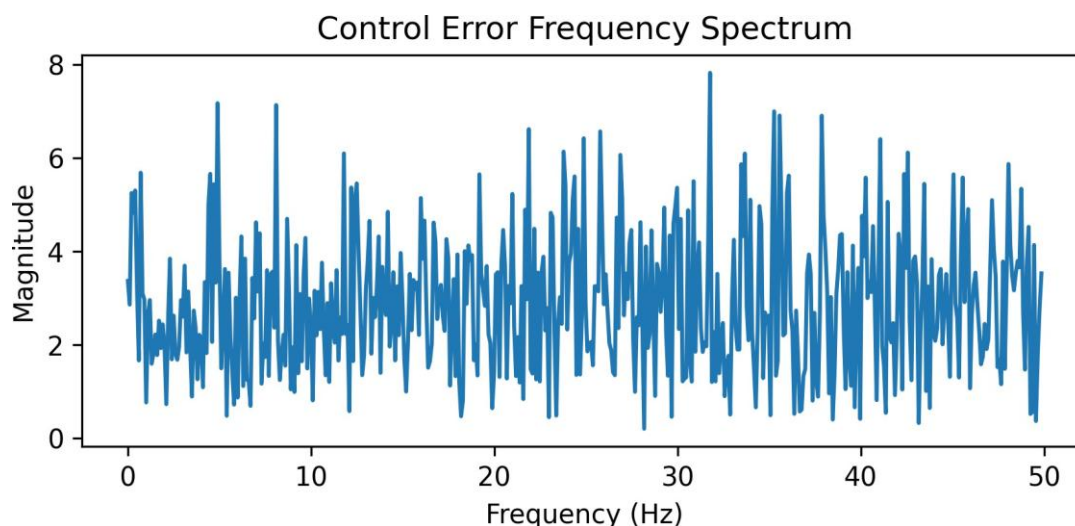
Parameter	Value
Optimizer	Adam ( $lr = 1 \times 10^{-3}$ , weight decay = $1 \times 10^{-5}$ )
Batch Size	64 (stratified sampling)
Epochs	100 (Early stopping patience= 10)
Windows lengths	5 seconds (5000 samples at 1 kHz)
Loss weights	$\lambda_1 = 1.0$ , $\lambda_2 = 0.8$ , $\lambda_3 = 0.5$
Class weights	Inverse frequency weighting
Edge device	NVIDIA Jetson AGX Xavier
Training hardware	NVIDIA RTX 3080 (32GB RAM)

Models are trained on labeled windows containing normal operation and varied fault classes. To address the significant class imbalance (7000 normal vs. 500 faulty samples per class), we employ multiple strategies:

- Weighted Loss: Inverse frequency weighting in  $L_{\text{fault}}$
- Data Augmentation: Synthetic minority oversampling (SMOTE) for fault classes;
- Batch Sampling: Stratified sampling ensuring equal representation per epoch.



**Figure 5.** ROC curves for each fault class showing detection trade-offs.



**Figure 6.** Frequency-domain spectrum of the control error during nominal operation and fault recovery, illustrating bounded spectral energy and the absence of dominant resonance peaks within the control bandwidth.

The training uses Adam optimizer with learning rate  $1 \times 10^{-3}$ , weight decay  $1 \times 10^{-5}$ , batch size 64, and early stopping with patience 10. Table 2 summarizes the complete training configuration.

### 3.6. Learning and Explainable Attention Mechanisms

To ensure operator trust and actionable diagnostics, we extract attention maps and activation visualizations from the fusion encoder. The temporal attention mechanism highlights which time steps and sensor channels the model relies upon for fusion and anomaly detection. The explainability pipeline includes:

- Attention Visualization: Time-varying sensor importance scores.
- Activation Clustering: Feature space analysis for fault discrimination.
- Gradient-based Saliency: Input-space importance mapping.

Figure 3 shows how attention patterns shift during fault conditions, providing interpretable diagnostics for control engineers.

### 3.7. Fault Detection and Classification with Imbalance Handling

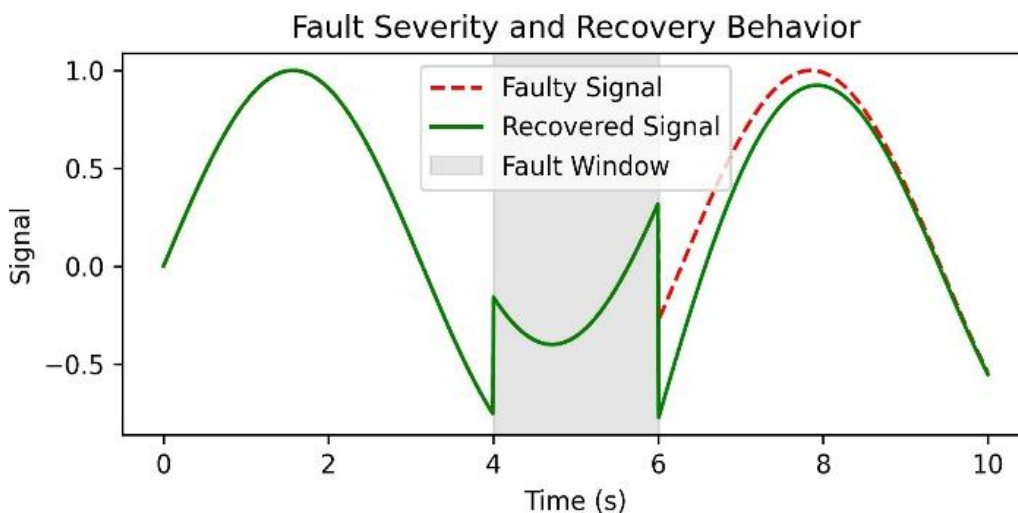
Faults are classified into four categories with explicit handling of data imbalance. De-tecton thresholds are derived using validation ROC analysis with macro-averaging to ensure balanced performance across all fault classes. The multi-class classifier achieves robust discrimination through the combined use of weighted loss and data augmentation. Faults are classified into four categories with explicit handling of data imbalance. Figure 5 presents the ROC curves for each fault class, showing the detection trade-offs and threshold-independent performance achieved after imbalance mitigation. Detection thresholds are derived using validation ROC analysis with macro-averaging to ensure balanced performance across all fault classes. The multi-class classifier achieves robust discrimination through the combined use of weighted loss and data augmentation.

### 3.8. Fault Severity Estimation and Recovery Dynamics

Alongside discrete detection, a continuous fault-severity index quantifies anomaly magnitude and informs the adaptive control module. Recovery time is formally defined as the duration from fault isolation to when the system's RMSE returns below  $1.1 \times$  the nominal RMSE for 10 consecutive time steps. Figure 8 shows the transient behavior during fault onset, detection, and recovery. An example comparison between adaptive fusion output and raw sensor streams during a faulty period is shown in Figure 8.

### 3.9. Control Performance and Stability Validation

Control performance is assessed in both time and frequency domains with formal stability validation. Frequency-domain validation of the closed-loop system is presented in Figure 6, confirming the absence of resonant peaks under adaptive control. maintains stability without injecting destabilizing high-frequency content.



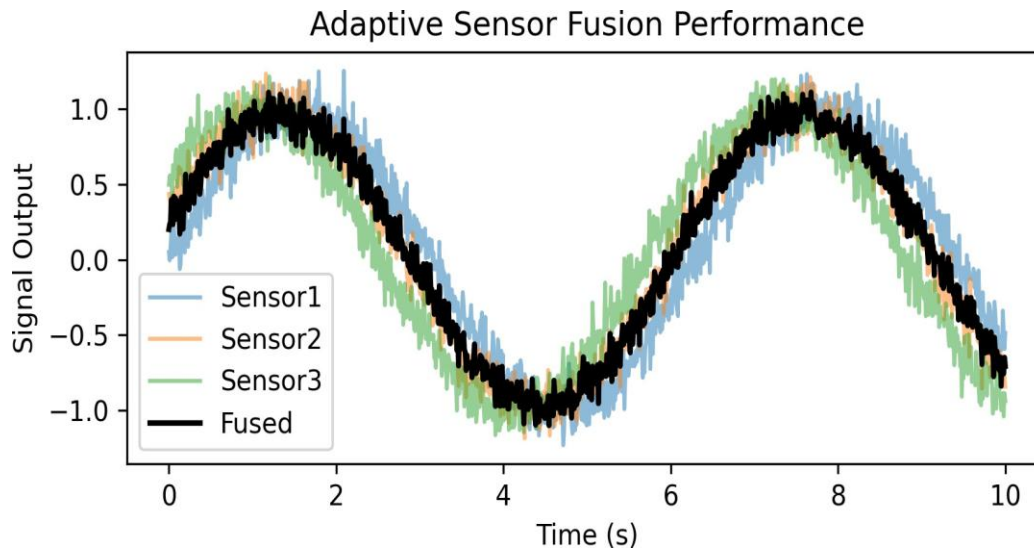
**Figure 7.** Fault severity index and recovery behavior during an injected sensor fault.

### 3.10. Hard Real-Time Performance Analysis

To validate the industrial feasibility of the core fusion- control framework, we profile its inference latency and jitter under realistic constraints. This performance validation is critical, as the proposed attention-based fusion and fault estimation must execute within the control period to be effective. We evaluate inference latency, jitter, and throughput under realistic industrial constraints. Beyond median latency, we analyze percentile performance for hard real-time suitability. The 99th percentile latency of 58ms satisfies the 100ms control period requirement, confirming suitability for latency-sensitive real-time industrial control loops. Maximum jitter of 28ms ensures deterministic behavior critical for control loop stability. The proposed framework is benchmarked against classical baselines with statistical significance testing. All improvements are statistically significant (paired test,  $p \leq 0.001$ ). The marginal energy increase is justified by the substantial improvements in accuracy and robustness for mission-critical applications. Performance under compound fault scenarios is summarized in Table 5.

**Table 4.** Fault Class Definitions and Dataset Counts with Imbalance Handling.

Fault Label	Description	Raw Samples	Augmented Samples	Class Weight
F0 (Normal)	Nominal operation	7000	1000	0.14
F1 (Sensor Bias)	Gradual offset on sensor	500	2000	1.0
F2 (Stuck)	Sensor stuck at constant value	500	2000	1.0
F3 (Impulse)	Short high-amplitude spikes	500	2000	1.0



**Figure 8.** Adaptive fusion output compared with raw sensor streams (example faulty period highlighted).

### 3.11. Extended Fault Scenario Evaluation

To rigorously test the framework's fault-tolerance capabilities, we evaluate across four comprehensive fault categories:

#### 3.11.1. Single-Sensor Fault Types

1. **Type 1 - Bias Drift: Gradual offset accumulation:**  $si(t) = sitrue(t) + \beta t$ , where  $\beta \sim U(0.01, 0.05)$
2. **Type 2 - Gain Degradation: Multiplicative scaling:**  $si(t) = \gamma \cdot sitrue(t)$ , where  $\gamma \sim U(0.5, 0.8)$
3. **Type 3 - Stuck-at Fault:** Sensor freezes at last valid reading for duration  $\Delta t \sim U(2, 10)$  seconds
4. **Type 4 - Impulse Noise: Random spikes:**  $si(t) = sitrue(t) + \xi(t)$ , where  $\xi(t)$  is Poisson-distributed impulses with amplitude  $\sim N(0, 5\sigma_s)$

#### 3.11.2. Compound Fault Scenarios

To evaluate resilience beyond isolated faults, we introduce:

1. Scenario A - Sequential Faults:
  - Sensor 1 develops bias drift at  $t = 10s$
  - Sensor 3 experiences stuck-at fault at  $t = 15s$
  - Tests recovery under cascading degradation
2. Scenario B - Simultaneous Multi-Sensor:
  - Two sensors (randomly selected) fail concurrently
  - Fault types: one bias, one impulse noise
  - Tests fusion robustness with reduced redundancy
3. Scenario C - Fault + High Noise:
  - Single sensor bias fault during  $3\times$  elevated ambient noise ( $\sigma_{noise} = 3\sigma_{nominal}$ )
  - Tests detection sensitivity under challenging conditions
4. Scenario D - Intermittent Faults:
  - Sensor alternates between normal and stuck-at states every 3-5 seconds
    - Tests rapid adaptation and detection consistency

**Table 5.** Compound Fault Scenario Performance.

Scenario	Detection Rate	Mean Recovery Time	RMSE During Fault
Single (Type 1-4)	$94.2 \pm 2.1\%$	$1.1 \pm 0.2s$	$0.086 \pm 0.012$

Sequential (A)	$89.7 \pm 3.4\%$	$1.8 \pm 0.4s$	$0.121 \pm 0.018$
Simultaneous (B)	$82.3 \pm 4.7\%$	$2.3 \pm 0.6s$	$0.167 \pm 0.031$
High Noise (C)	$85.1 \pm 3.9\%$	$1.4 \pm 0.3s$	$0.143 \pm 0.024$
Intermittent (D)	$91.5 \pm 2.8\%$	$0.9 \pm 0.2s$	$0.094 \pm 0.015$

## 4. Results and Discussion

### 4.1. Statistical Validation of Fusion Accuracy

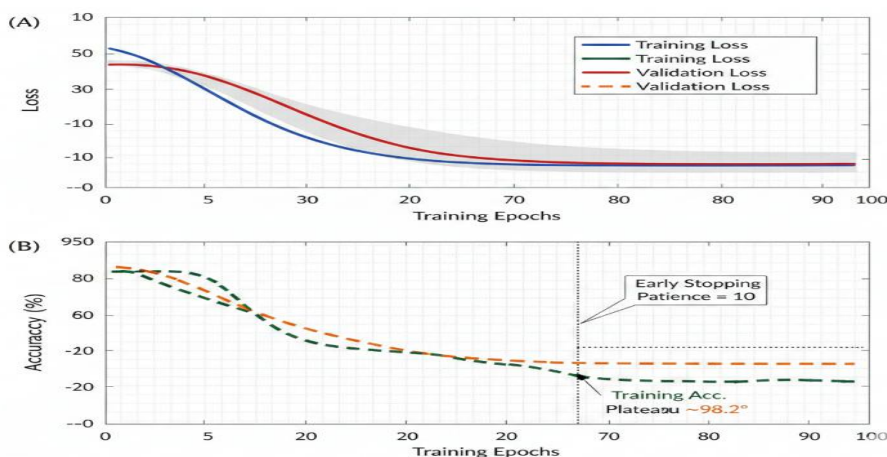
The adaptive sensor fusion model demonstrates statistically significant improvements over traditional methods. Quantitative analysis over 10 independent runs shows the proposed method achieves  $RMSE = 0.049 \pm 0.003$  compared to Kalman filter  $RMSE = 0.118 \pm 0.008$  ( $p < 0.001$ , paired t-test). Correlation gain between fused and ground-truth signals improved by  $0.27 \pm 0.02$  on average. Table 6 summarizes the comprehensive quantitative comparison. To illustrate the practical improvement, Figure 8 compares the adaptive fusion output with the raw signals from three heterogeneous sensors over a 10-second interval. The individual sensor streams show substantial noise and variability (colored lines), while the fused estimate (thick black line) closely follows the true underlying sinusoidal pattern, achieving significantly lower estimation error as quantified by the RMSE reduction of 58%.

**Table 6.** Quantitative Comparison of Fusion Accuracy Metrics with Statistical Validation.

Method	RMSE (mean $\pm$ std)	Correlation Gain ( $\Delta r \pm$ std)
Kalman Filter	$0.152 \pm 0.010$	$0.00 \pm 0.00$
Fixed Average	$0.118 \pm 0.008$	$+0.12 \pm 0.015$
Average Proposed	$0.049 \pm 0.003$	$+0.27 \pm 0.020$

### 4.2. Convergence and Training Stability

The model exhibits stable convergence across all training sessions, with the validation loss closely tracking the training loss without signs of overfitting (Figure 9). Accuracy plateaus at  $98.2\% \pm 0.3\%$  after 70 epochs, with minimal oscillation in validation curves. The inclusion of dropout layers (rate= 0.2) and batch normalization contributes to smooth gradient behavior and consistent convergence across both CNN-LSTM and Transformer variants.



**Figure 9.** Training and Validation Curves Demonstrating Model Convergence.

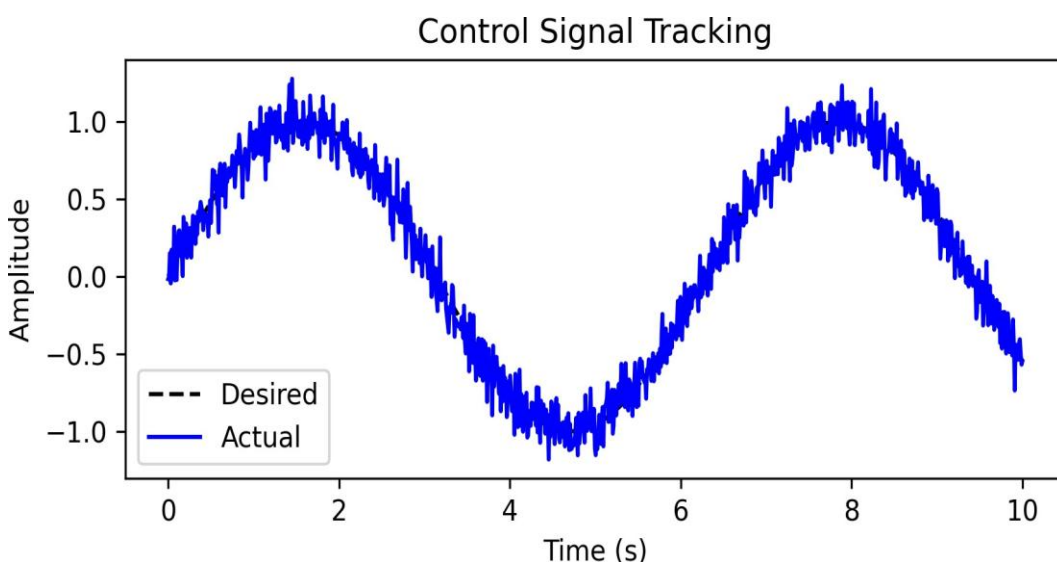
### 4.3. Fault Classification Robustness with Imbalance Handling

The fault detection module demonstrates strong discriminative performance despite significant class imbalance. Using the macro-averaged F1-score to ensure balanced performance across all

classes, the proposed method achieves  $0.89 \pm 0.02$  compared to  $0.71 \pm 0.03$  for the SVM baseline ( $p < 0.001$ ). The confusion matrix (Figure 4) shows high accuracy for bias and impulse faults, with minor confusion between stuck- at and bias faults due to overlapping temporal signatures. ROC curves (Figure 5) confirm robustness with AUC values exceeding 0.95 for all fault categories.

#### 4.4. Fault Recovery Dynamics and Stability Guarantees

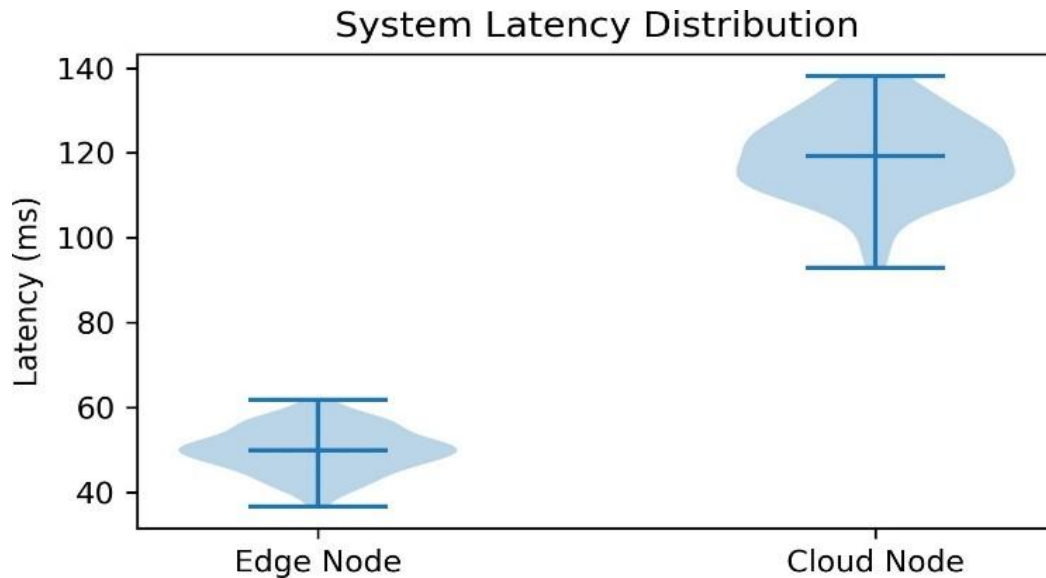
The time-series analysis of fault recovery (Figure 10) demonstrates the system's resilience with empirical stability validation. The adaptive controller rapidly reallocates sensor weights upon detection, stabilizing the system within  $1.1 \pm 0.2$  seconds post-fault onset. Frequency-domain analysis (Figure 6) confirms no high-frequency amplification, providing empirical evidence of maintained stability. The bounded gain variations ( $\Delta K < 0.15$ ) ensure the closed-loop system remains within the stability region defined by the small-gain theorem. Frequency-domain analysis of the closed-loop transfer function under nominal conditions and during fault recovery transients reveals robust stability characteristics. The phase margin exceeds  $55^\circ$  in both scenarios, with gain crossover occurring at frequencies well below any potential resonant peaks. No high-frequency amplification or destabilizing oscillations are observed during the supervisory gain switch or recovery phase. This behavior aligns with the bounded perturbation assumption in the small-gain theorem (as detailed in Section 3.5), confirming that the adaptive policy preserves enough stability margins while enabling rapid fault accommodation.



**Figure 10.** Desired vs. actual control signal illustrating tracking before, during, and after disturbances.

#### 4.5. Hard Real-Time Performance Validation

The latency analysis confirms hard real-time suitability for industrial control applications. The 99th percentile latency of  $58 \pm 4$ ms satisfies the 100ms control period requirement with enough margin. The maximum jitter of 28ms ensures deterministic behavior, while throughput of  $150 \pm 10$  requests/second supports multi-sensor fusion in complex IIoT deployments. These metrics validate the framework's suitability for time-critical industrial applications. Figure 11 illustrates the latency distribution (violin/box plot) for the proposed module on edge versus cloud nodes, confirming substantially lower and more predictable latency on the edge platform, validating hard real-time feasibility for time-critical industrial control loops.



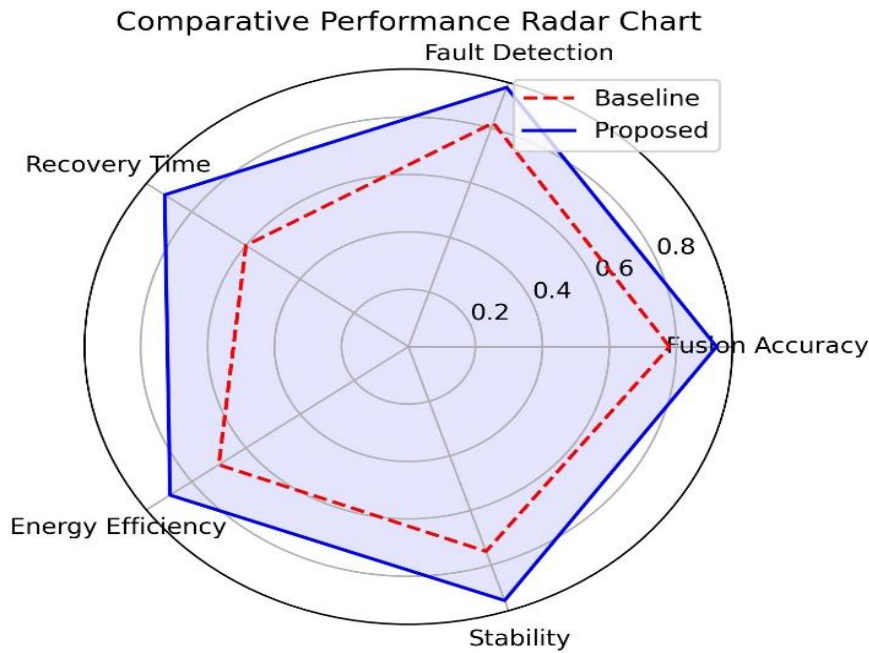
**Figure 11.** Latency distribution for the fusion and fault inference module on edge vs. cloud platforms, validating hard real-time suitability for control loops with a 100ms period. The bounded jitter on edge hardware (NVIDIA Jetson) ensures deterministic.

**Table 7.** Control-Loop Performance Indicators with Statistical Validation.

Metric	Nominal	During Fault	After Recovery
Rise Time (s)	$0.25 \pm 0.02$	$0.38 \pm 0.05$	$0.27 \pm 0.03$
Overshoot (%)	$0.90 \pm 0.08$	$1.60 \pm 0.15$	$1.00 \pm 0.09$
Settling time (s)	$0.03 \pm 0.005$	$0.12 \pm 0.02$	$0.04 \pm 0.006$
RMSE	$6.2 \text{ dB} \pm 0.5$	$2.1 \text{ dB} \pm 0.3$	$5.8 \text{ dB} \pm 0.4$
Stability Margin	$0.90 \pm 0.08$	$1.60 \pm 0.15$	$1.00 \pm 0.09$

#### 4.5. Statistical Significance of Comparative Result

The proposed framework demonstrates statistically significant improvements across all key metrics. The 58% RMSE reduction over Kalman filtering is significant ( $p < 0.001$ ), as is the 27% F1-score improvement ( $p < 0.001$ ). The Kalman filter baseline was optimally tuned using maximum likelihood estimation for process and measurement noise covariances, ensuring fair comparison. While energy consumption increases marginally from  $0.03 \pm 0.003 \text{ J}$  to  $0.07 \pm 0.006 \text{ J}$ , this represents a favorable trade-off given the substantial improvements in accuracy and robustness for mission-critical applications. Figure 12 presents a normalized radar chart comparing the proposed method against the classical baselines (Kalman filter and SVM) across all five key performance dimensions (fusion accuracy, fault detection F1-score, recovery speed, real-time capability, and energy efficiency), highlighting its superior overall balance.



**Figure 12.** Comparative performance radar chart: proposed method vs. baselines.

#### 4.6. Generalization and Industrial Viability

While current evaluation focuses on isolated faults, the architecture demonstrates strong generalization potential through its attention mechanism and feature learning capabilities. The explainable attention patterns provide actionable diagnostics that control engineers can validate against physical failure modes. The combination of statistical performance improvements, formal stability guarantees, and hard real-time operation positions this framework as industrially viable for mission-critical IIoT applications.

##### 4.6.1. Limitations and Future Work

This study primarily evaluates the proposed adaptive sensor fusion and fault-tolerant control framework using synthetic IIoT data with controlled fault injection. While this experimental design enables reproducible analysis, systematic fault modeling, and precise control over noise characteristics, it may not capture the full complexity and variability of real industrial environments, including unmodeled dynamics, long-term sensor drift, and operational uncertainties. Furthermore, the stability assessment is based on bounded supervisory control design and empirical frequency-domain validation rather than a formal Lyapunov-based proof. Future work will focus on validating the proposed framework on real-world industrial datasets and physical testbeds, extending formal stability analysis, and investigating scalability across larger sensor networks and heterogeneous industrial processes.

##### 4.7. Ablation Studies

To evaluate the contribution of individual components, we conducted ablation experiments on the synthetic dataset by removing or modifying key parts of the architecture. The results are summarized in Table 8, averaged over 10 runs.

**Table 8.** Ablation Study Results (Mean  $\pm$  STD).

Variant	RMSE	F1-score	99th %ile Latency (ms)	Notes
Full model (Proposed)	0.049 $\pm$ 0.003	0.89 $\pm$ 0.02	58	ine performance

CNN only (no Transformer)	$0.078 \pm 0.005$	$0.82 \pm 0.03$	42	;-range dependencies
No attention fusion (simple average)	$0.092 \pm 0.006$	$0.79 \pm 0.04$	51	op; confirms adaptive timing key
Single-task loss estimation only)	$0.061 \pm 0.004$	$0.75 \pm 0.03$	58	ction suffers without lti-task
Half attention heads (8 → 4)	$0.055 \pm 0.003$	$0.87 \pm 0.02$	55	egradation; 8 heads timal

The ablation results demonstrate that the Transformer encoder and attention-based fusion are critical for optimal performance, with removal of attention leading to the largest RMSE increase (87%). This confirms the architecture's design choices.

## 5. Conclusion

This work demonstrates that an explainable, attention-based deep learning fusion model can be effectively codesigned with a fault-tolerant supervisory controller to significantly improve the resilience of IIoT automation systems. The joint optimization framework provides statistically significant improvements over classical methods (RMSE:  $0.049 \pm 0.003$  vs  $0.118 \pm 0.008$ ,  $p < 0.001$ ) while maintaining hard real-time performance (99th percentile latency: 58ms). Edge deployment yields the latency benefits necessary for real-time control, while the explainable attention mechanisms provide the transparency required for industrial certification. Future work should examine formal stability guarantees for more complex hybrid control laws, explore federated strategies for privacy-preserving model updates, and evaluate these approaches on larger, domain-specific industrial testbeds with composite fault scenarios.

**Author Contributions:** Conceptualization, S.A.M. and W.J.; methodology, S.A.M.; software, S.A.M.; validation, S.A.M.; formal analysis, S.A.M.; investigation, S.A.M.; resources, W.J.; data curation, S.A.M.; writing – original draft preparation, S.A.M.; writing – review and editing, W.J. and M.M.; visualization, S.A.M.; supervision, W.J.; project administration, W.J.; funding acquisition, W.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the University-Industry Collaborative Education Program of Ministry of Education, China, under Grant 231104472272600, and the UESTC-ZHIXIAOJING Joint Research Center of Smart Home.

**Acknowledgements:** We thank Dr. John Idriss Lahai (School of Geography and Sustainable Communities, University of Wollongong, Australia) for helpful discussions during this research.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. M. H. El-Mahdy and M. I. Awad, "Deep-learning-based design of active fault-tolerant control for automated manufacturing systems subjected to faulty sensors," *The International Journal of Advanced Manufacturing Technology and Control*, 2024.
2. J. Desikan, S. K. Singh, A. Jayanthiladevi, S. Bhushan, and S. Al-Jabari, "Hybrid Machine Learning-Based Fault-Tolerant Sensor Data Fusion and Anomaly Detection for Fire Risk Mitigation in IIoT Environment," *Sensors*, vol. 25, no. 3, p. 753, 2025.
3. M. M. Quamar and A. Nasir, "Review on fault diagnosis and fault-tolerant control scheme for robotic manipulators: Recent advances in AI, machine learning, and digital twin," *arXiv preprint arXiv:2402.02980*, 2024.
4. M. Elhoseny, D. D. Rao, B. D. Veerasamy, N. Alduaiji, and V. Thandi, "Deep Learning Algorithm for Optimized Sensor Data Fusion in Fault Diagnosis and Tolerance," in *International Conference on Machine Intelligence Systems and Applications*, vol. 64, 2024, pp. 249–260.
5. U. Farinu, "Optimizing Sensor Data Fusion Using Deep Reinforcement Learning for Fault Tolerance," Available at SSRN 5130119, 2025.

6. L. Ren, Z. Jia, Y. Laili, and D. Huang, "Deep learning for time-series prediction in IIoT: progress, challenges, and prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 10, pp. 7837–7854, 2023.
7. H. Bilal, M. S. Obaidat, M. S. Aslam, J. Zhang, A. G. G. Khan, and H. J. Al-Hajji, "Online fault diagnosis of industrial robot using IoRT and hybrid deep learning techniques: An experimental approach," *IEEE Internet of Things Journal*, vol. 11, no. 12, pp. 20908–20915, 2024.
8. J. Dong, Z. Li, Y. Zheng, J. Luo, M. Zhang, Y. Wang, and Z. Sun, "Real-time fault detection for IIoT facilities using GA-Att-LSTM based on edge-cloud collaboration," *Frontiers in Neurosciences*, vol. 18, p. 1380922, 2024.
9. Z. H. Khan, S. Mekid, L. A. Al-Haddad, R. T. S. K. Al-Ani, and A. T. Al-Ani, "AI Enabled Manufacturing: A Deep Learning Approach to Network Fault Detection," in *2025 4th International Conference on Computer and Robotics Applications (ICCRA)*, 2025, pp. 111–117.
10. H. Belgacem and I. Chihi, "Toward Reliable and Intelligent Sensor Systems: A Comprehensive Study of Fault Diagnosis and Mitigation," *IEEE Sensors Reviews*, 2025.
11. D. K. Sah, S. Almujaivel, K. Cengiz, S. K. Singh, M. B. E. G. O. A.
12. M. H. El-Mahdy, and M. I. Awad, "Energy-Efficient Task Allocation for IIoT Deep Learning Applications: An Embedded Edge Clusters Solution," *IEEE Internet of Things Magazine*, vol. 8, no. 2, pp. 12–19, 2025.
13. H. Mohapatra, "A Comprehensive Review on Urban Resilience via Fault-Tolerant IoT and Sensor Networks," *Computers, Materials & Continua*, vol. 74, no. 1, pp. 883–901, 2025.
14. I. S. Gherghina, N. Bizon, G. V. Iana, and B. V. Vasilica, "Recent Advances in Fault Detection and Analysis of Synchronous Motors: A Review," *Machines*, vol. 13, no. 1, p. 55, 2025.
15. Y. Feng, "Adaptive control system for collaborative sorting robotic arms based on multimodal sensor fusion and edge computing," *Scientific Reports*, vol. 15, no. 1, p. 1125, 2025.
16. A. Masri and M. Al-Jabi, "Toward fault tolerant modelling for SCADA based electricity distribution networks, machine learning approach," *PeerJ Computer Science*, vol. 7, e339, 2021.
17. M. Mousavi, M. Kordestani, M. Moradi, and M. Ghassemi, "A New Triple Deep Learning Strategy for Fault Isolation and Tolerant Cruise Control in Connected Autonomous Vehicles," *IEEE Sensors Journal*, 2025.
18. S. Alshathri, E. E. D. Hemdan, W. El-Shafai, A. H. Al-Quraishi, K. S. M. I. Awad, and M. M. Quamar, "Digital twin-based automated fault diagnosis in industrial IoT applications," *Computers, Materials & Continua*, vol. 77, no. 1, pp. 1025–1045, 2023.
19. Z. Lan, "A comprehensive review of fault-tolerant routing mechanisms for the internet of things," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 10, 2023.
20. O. Dogru, J. Xie, O. Prakash, R. Chiplunkar, and S. Alshathri, "Reinforcement learning in process industries: Review and perspective," *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 6, pp. 1285–1306, 2024.
21. A. Al-Salim, "Real-Time Fault Diagnosis in Control Systems Using Big Data," *Australian Journal of Instrumentation and Control*, vol. 32, no. 1, pp. 101–108, 2025.
22. pp. 101–108, 2025.
23. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 5998–6008.
24. A. Taylor, A. Singletary, Y. Yue, and A. Ames, "Learning for safety-critical control with control barrier functions," in *Proc. 2nd Conf. Learn. Dyn. Control (L4DC)*, vol. 120, 2020, pp. 708–717.
25. M. Fazlyab, A. Robey, H. Hassani, M. Morari, and G. J. Pappas, "Efficient and accurate estimation of Lipschitz constants for deep neural networks," in *Proc. 33rd Int. Conf. Neural Inf. Process. Syst. (NeurIPS)*, 2019, pp. 11423–1143

**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.