aerospace

MDPI

Article

Real-Time Anomaly Detection for Civil Aviation VHF Communications Using Learnable Kernels and Conditional GANs

Junyi Zhai [1], Gang Sun [1], Zhengqiang Li [2], Quanxin Cao [2] and Yufeng Huang [1,*]

[1] Department of Aeronautics and Astronautics, Fudan University, Shanghai 200433, China; zhaijunyi@comac.cc (J.Z.); gang_sun@fudan.edu.cn (G.S.)
[2] Shanghai Aircraft Design and Research Institute, Shanghai 201210, China; lizhengqiang@comac.cc (Z.L.); caoquanxin@comac.cc (Q.C.)
* Correspondence: yufenghuang@fudan.edu.cn

Abstract

Civil aviation VHF communication is safety-critical, yet operational links are routinely disturbed by atmospheric effects, aging hardware, and electromagnetic interference. The resulting anomalies are typically weak, intermittent, and extremely rare, which makes real-time detection difficult under strong temporal dependence and severe class imbalance. We propose an end-to-end framework that couples (i) a learnable kernel projection for adaptive nonlinear feature extraction, (ii) a differentiable relevance–redundancy objective for feature refinement, and (iii) conditional temporal generation to augment minority anomaly patterns. A lightweight CNN–LSTM head is used for streaming inference. Training uses a mixture of operational anomalies and simulated degradation scenarios, while evaluation is conducted using operational data only. Experiments on 1.2 million VHF frames collected from real flight operations and ground station monitoring achieve an F1-score of 0.947, ROC-AUC of 0.972, and PR-AUC of 0.968, with an average inference latency of 34.7 ms.

Keywords: VHF communication; anomaly detection; learnable kernel; differentiable MRMR; conditional GAN; temporal generation; hybrid deep learning; PHM

1. Introduction

Very High Frequency (VHF) communication continues to play a central role in civil aviation voice and data exchange within line-of-sight coverage. It is used for routine air traffic control coordination, flight operations, and emergency communication, so communication reliability is tightly linked to operational safety and efficiency [1]. Despite its long history, maintaining stable performance under real operating conditions is still challenging.

In operational environments, VHF links are exposed to many sources of disturbance. Atmospheric effects, stochastic propagation, hardware aging, multipath propagation, and electromagnetic interference can all affect signal quality. These factors rarely lead to immediate failure; instead, they often appear as intermittent degradations such as packet loss, rising bit-error rates, signal fading, and latency variation. Although each effect may seem minor on its own, their combined impact can gradually reduce communication reliability when early-warning mechanisms are absent [2–4].

Figure 1 presents a simplified view of a civil aviation VHF communication system. Aircraft communicate with ground stations through line-of-sight VHF channels, and the resulting communication records are forwarded to air traffic control centers for monitoring

Check for updates

Academic Editor: Miroslav Kelemen

Received: 12 January 2026
Revised: 20 February 2026
Accepted: 10 March 2026
Published: 13 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

and decision making. To extend coverage, modern aviation infrastructures increasingly rely on relay links and satellite-assisted communication, while these technologies improve connectivity, they also introduce additional heterogeneity and uncertainty, which increases the need for continuous and intelligent monitoring.

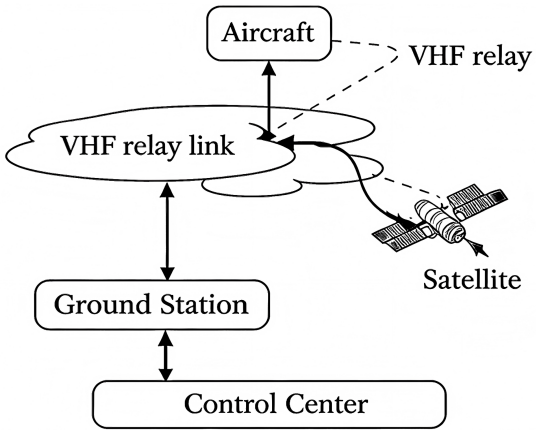


Figure 1. Simplified architecture of a typical civil aviation VHF communication system.

Traditional monitoring solutions such as Built-In Test Equipment (BITE), rule-based diagnostics, and periodic inspection provide only limited reliability assurance. These approaches are mainly designed to detect clear failures rather than subtle degradations in continuous communication streams. Another challenge is the severe class imbalance in operational monitoring data, where anomaly samples typically account for less than 1% of observations [5]. This combination of rarity, temporal dependence, and nonlinear feature interaction makes conventional rule-based and supervised methods difficult to apply effectively.

Deep learning has shown strong performance in wireless and network anomaly detection. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid spatiotemporal models can learn complex communication patterns directly from data. However, applying these advances to aeronautical VHF monitoring is not straightforward. Real-time latency constraints, safety-critical requirements, and the need for stable and interpretable behavior introduce additional design constraints that are often overlooked in general wireless scenarios [6–9].

Prognostics and Health Management (PHM) has improved the reliability of aerospace subsystems such as engines, avionics, and sensors [10,11]. Communication subsystems, however, have received comparatively less attention. Unlike mechanical components that often show gradual wear, communication channels are dominated by stochastic propagation and rapidly changing operating conditions. Effective VHF monitoring must therefore handle strong nonlinearity, extreme anomaly sparsity, and strict real-time requirements simultaneously [2,3].

Early communication anomaly detection mainly relied on statistical analysis and rule-based indicators derived from signal-quality metrics such as SNR, BER, and RSSI. Although interpretable and easy to deploy, these approaches typically assume stationary channel conditions and struggle to capture nonlinear and time-varying degradation patterns [2,3]. Classical machine learning methods, including Support Vector Machines and Random Forests, were later introduced to improve detection performance. These models often rely on hand-crafted features and analyze frames independently, which limits their ability to capture long-term temporal dependencies [9].

Recent work has explored deep learning for wireless anomaly detection. CNN-based models learn hierarchical signal representations, while recurrent architectures such as

LSTMs capture sequential communication behavior [7]. Hybrid architectures further combine convolutional feature extraction with sequence modeling to improve performance [12]. Despite these advances, many studies still focus mainly on detection accuracy and pay less attention to feature redundancy, extreme class imbalance, and real-time deployment requirements in safety-critical aviation settings.

In parallel, nonlinear representation learning and generative modeling have attracted growing attention. Kernel-based techniques such as KPCA provide effective nonlinear feature extraction but rely on fixed kernels that may become suboptimal under nonstationary channel conditions [13,14]. Mutual-information-based feature selection, particularly the MRMR criterion, has been widely used to reduce redundancy [15,16]. Generative Adversarial Networks (GANs) have also been used to address class imbalance through data augmentation; however, many existing approaches focus on frame-level generation and do not explicitly model temporal structure or communication-state conditions [17–19].

These limitations motivate the development of an integrated framework that addresses nonlinear representation learning, redundancy-aware feature selection, imbalance mitigation, and real-time sequential detection.

In this work, we propose an end-to-end anomaly detection framework that combines learnable nonlinear representation, differentiable redundancy-aware feature shaping, and condition-consistent temporal augmentation. A Learnable Kernel–Redundancy Projection (LKRP) module replaces fixed kernel mappings and discrete feature selection with a parametric kernel and a differentiable relevance–redundancy objective, enabling task-adaptive feature extraction for VHF communication signals [13]. A conditional temporal Wasserstein GAN is introduced to synthesize anomaly sequences using communication-state variables such as SNR, RSSI, and BER, improving minority anomaly coverage while preserving temporal coherence [18,19]. The resulting features are analyzed by a lightweight CNN–LSTM detector designed to satisfy real-time avionics latency requirements.

Civil aviation VHF communication operates under strict safety and regulatory requirements and is mainly used for pilot–ATC voice coordination. Communication quality varies with flight phase, traffic density, terrain shielding, and sector handovers. In this setting, degraded links may affect situational awareness and workload. These characteristics motivate the monitoring strategy adopted in this work.

The main contributions of this work can be summarized as follows:

- We propose a Learnable Kernel–Redundancy Projection module to obtain task-adaptive nonlinear representations for VHF communication monitoring.
- We develop a conditional temporal WGAN to generate anomaly sequences conditioned on communication states, which alleviates the extreme class imbalance.
- We design a lightweight CNN–LSTM detector that supports real-time inference under strict latency constraints.
- We validate the proposed framework on a large-scale operational VHF dataset and demonstrate improved detection performance and practical deployment feasibility.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the proposed methodology. Section 4 describes the dataset and experimental setup. Section 5 reports experimental results, and Section 6 concludes the paper.

2. Related Work

Research on anomaly detection for aviation VHF communication draws from several areas, including PHM, wireless signal monitoring, nonlinear representation learning, and generative modeling. Compared with terrestrial wireless networks, aeronautical communication operates under stricter safety, coverage, and reliability constraints. Rather than treating these research directions separately, this section reviews prior work from the per-

spective of complete anomaly-detection pipelines and identifies the remaining challenges for aviation VHF monitoring.

2.1. Discriminative Approaches for Communication Anomaly Detection

PHM has become an important enabler of reliability improvement and condition-based maintenance in aerospace systems. A large body of work has demonstrated its value for engines, avionics, sensors, and structural components, where degradation often evolves gradually and can be analyzed using historical trends [10,11,20,21].

Communication subsystems present a different set of challenges. VHF links are strongly affected by environmental variability and stochastic propagation, resulting in degradation patterns that are intermittent and difficult to model. As a result, monitoring efforts have traditionally focused on hardware diagnostics or coarse channel-quality indicators rather than frame-level anomaly detection in continuous communication streams [1,2,22]. Related work in aerospace and industrial PHM highlights the increasing role of data-driven monitoring in complex engineering systems [20,21,23,24].

Early work on communication anomaly detection relied on statistical thresholds, hypothesis testing, and classical machine learning methods such as Support Vector Machines and Random Forests. Although these approaches are interpretable and easy to deploy, they typically treat samples independently and therefore struggle to capture nonlinear temporal dependencies.

More recently, deep learning has improved anomaly detection in wireless and networked environments [25,26]. CNN-based models learn hierarchical signal representations, while recurrent architectures such as LSTMs capture temporal behavior in communication sequences [7]. Hybrid spatiotemporal models combine convolutional and recurrent components to further enhance performance [12]. Related work has also explored spectrum monitoring, software-defined networking, and model-fusion strategies for large-scale network anomaly detection [25,27,28].

Despite these advances, many discriminative approaches assume that datasets are relatively balanced and they rely on fixed feature representations. In aviation VHF monitoring, anomalies are extremely rare and channel conditions are highly nonstationary, which limits the robustness of purely discriminative models.

2.2. Kernel-Based Representation Learning and Feature Selection

Kernel Principal Component Analysis (KPCA) has long been used for nonlinear feature extraction in monitoring and fault diagnosis [13,14]. By mapping observations into a reproducing kernel Hilbert space, KPCA enables linear analysis of complex nonlinear structures. However, classical KPCA relies on predefined kernels and fixed projections, which may become suboptimal as data distributions evolve.

Recent studies have investigated learnable kernel parameterizations that adapt to specific tasks and datasets [13]. More broadly, advances in neural representation learning emphasize the importance of adaptive feature extraction for complex monitoring data. Sparse representation methods such as matching-pursuit networks can learn interpretable time–frequency representations for fault diagnosis [29], while attention-guided graph models capture structural dependencies and improve reliability prediction in PHM applications [30]. These studies highlight the need for flexible feature representations in complex monitoring scenarios.

High-dimensional monitoring data also contain redundant or weakly informative features. Mutual-information-based feature selection, particularly the minimum-redundancy–maximum-relevance (MRMR) criterion, is widely used to address this issue [15,16]. However, MRMR is usually implemented through greedy search and offline

estimation, making it difficult to integrate into gradient-based end-to-end learning frameworks. Consequently, representation learning and feature selection are rarely optimized jointly with the detection model.

2.3. Generative Models for Imbalanced Anomaly Detection

Class imbalance remains a major challenge in anomaly detection. Generative Adversarial Networks (GANs) have therefore been widely explored as a way to augment minority classes [17]. Wasserstein GANs with gradient penalty (WGAN-GP) improve training stability and have been successfully applied in fault diagnosis and anomaly detection [18,19,31].

Many existing approaches focus on static or frame-level data. Temporal dependencies are often overlooked, and communication-state variables such as SNR, RSSI, and BER are rarely incorporated into the generation process, which limits the realism and diversity of synthetic samples [32–35].

Recent studies have extended GAN-based anomaly detection beyond traditional fault-diagnosis scenarios. GAN variants have been applied to industrial control systems using dual auto-encoder architectures to capture complex system dynamics [33]. Temporal GAN models such as TadGAN highlight the importance of sequence-level generation for time series anomaly detection [34]. In wireless and networking contexts, conditional GAN-based augmentation has been investigated to improve robustness under distribution shifts and limited anomaly samples [35]. Together, these studies show that incorporating temporal structure and contextual information is important for generating realistic minority anomalies.

2.4. Research Gap

Although substantial progress has been made in communication and PHM anomaly detection, several challenges remain for real-time VHF monitoring. Discriminative models often rely on fixed representations and struggle under extreme class imbalance. Kernel projection and feature selection are usually performed offline and cannot be optimized jointly with the detector. Current generative approaches also rarely combine temporal modeling with communication-state conditioning.

These limitations motivate the development of a unified framework that combines learnable kernel-based representation learning, differentiable redundancy-aware feature shaping, and conditional temporal generative augmentation for real-time VHF anomaly detection.

3. Methodology

3.1. Framework Overview

This work develops a unified framework for real-time anomaly detection in VHF communication signals. The framework integrates nonlinear feature projection, redundancy-aware representation learning, condition-aware temporal data augmentation, and lightweight sequential detection within a single processing pipeline.

Let $\mathbf{X} = \{x_1, \ldots, x_N\} \in \mathbb{R}^{N \times d}$ denote the raw VHF feature set and $\mathbf{C} = \{c_1, \ldots, c_N\}$ the corresponding labels, where $c_i \in \{0, 1\}$ indicates normal and anomalous states [2,5]. Each frame x_i is associated with a communication condition vector s_i (e.g., SNR, RSSI, BER, and PER), which provides contextual information used during condition-aware augmentation.

The processing pipeline is illustrated in Figure 2. The workflow begins with preprocessing and standardization of raw VHF communication frames. The normalized features are then mapped into a nonlinear representation space using the Learnable Kernel–Redundancy Projection (LKRP) module [13]. To address severe class imbalance, a conditional temporal WGAN is trained to generate anomaly-like sequences in the learned

feature space [18,19,31,36]. A sliding window CNN–LSTM detector then produces real-time anomaly scores for continuous monitoring [12].

Overall Framework of the Proposed LKRP–CT-WGAN-Enhanced Real-Time Anomaly Detection System

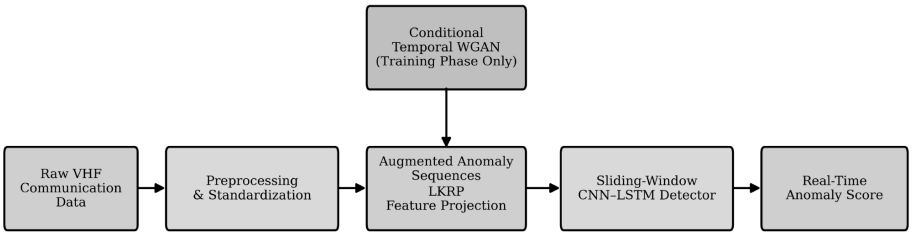


Figure 2. Overview of the LKRP–CT–WGAN–CNN–LSTM pipeline for real-time anomaly detection in civil aviation VHF communication systems. The framework combines learnable nonlinear feature projection, condition-aware temporal augmentation, and hybrid spatiotemporal detection under severe class imbalance and strict avionics latency requirements.

The LKRP module produces compact feature representations that preserve nonlinear structure under nonstationary channel conditions. During training, CT–WGAN increases the diversity of anomaly samples in a condition-consistent manner. The CNN–LSTM detector operates on the resulting feature sequences and outputs anomaly scores suitable for continuous onboard and ground monitoring.

For clarity, we summarize the overall learning objective of the framework as a single loss that collects the contributions of LKRP, CT–WGAN, and the CNN–LSTM detector:

$$\mathcal{L}_{total} = \mathcal{L}_{clf} + \alpha\mathcal{L}_{CT} + \beta\mathcal{L}_{LK} + \gamma\mathcal{L}_{DMR}. \quad (1)$$

Here, $\mathcal{L}_{clf}$ denotes the CNN–LSTM classification loss [12], $\mathcal{L}_{CT}$ represents the conditional temporal WGAN adversarial loss [18,19,31,36], $\mathcal{L}_{LK}$ regularizes structure preservation in the learnable kernel projection [13], and $\mathcal{L}_{DMR}$ encourages high relevance and low redundancy via a differentiable minimum-redundancy–maximum-relevance (MRMR) objective, hereafter referred to as differentiable MRMR (DMR) [15,16].

Training Strategy

Equation (1) summarizes the overall objectives of the framework. Although the losses are jointly defined, the modules are trained sequentially in practice to improve stability and convergence.

First, LKRP and DMR are pretrained using $\mathcal{L}_{LK} + \gamma\mathcal{L}_{DMR}$. Second, CT–WGAN is trained on LKRP embeddings of anomaly sequences. Finally, the CNN–LSTM detector is trained on the augmented dataset and LKRP is jointly fine-tuned with the detector using $\mathcal{L}_{clf}$. During inference, the CT–WGAN module is removed and only LKRP projection and the CNN–LSTM detector are executed.

3.2. Data Acquisition and Preprocessing

Raw VHF communication data are collected from aircraft operational flight logs and ground station monitoring records, supplemented by controlled channel-degradation scenarios designed to emulate representative real-world disturbances. For each communication frame x_i, a set of key radio frequency and link quality indicators is extracted, including signal-to-noise ratio (SNR), bit error rate (BER), received signal strength indicator (RSSI), packet error rate (PER), retransmission count, and transmission latency. These quan-

tities form the raw feature vector $x_i \in \mathbb{R}^d$ and the condition vector s_i used for conditional generation [2,3].

3.2.1. Dataset Clarification

The training data combine operational anomalies with simulated degradation scenarios in order to increase the coverage of rare abnormal patterns. In contrast, the test set contains only operational anomalies so that the reported results reflect performance under real-world conditions. Simulated data are used only for training and are not used in the final evaluation.

3.2.2. Operational Data Context

The dataset originates from routine civil aviation VHF communication operations involving short- and medium-haul aircraft. The recordings cover multiple flight phases, including climb, cruise, and descent, and were obtained from different airspace regions with varying traffic density and propagation conditions. These variations introduce realistic channel nonstationarity and environmental diversity into the monitoring data.

3.2.3. Feature Standardization

Let $x_{i,j}$ denote the j-th feature of frame x_i. Each feature dimension is standardized to zero mean and unit variance:

$$x_{i,j}^{\text{norm}} = \frac{x_{i,j} - \mu_j}{\sigma_j}, \qquad \mu_j = \frac{1}{N}\sum_{i=1}^{N} x_{i,j}, \qquad \sigma_j^2 = \frac{1}{N}\sum_{i=1}^{N}(x_{i,j} - \mu_j)^2. \tag{2}$$

In matrix form, the normalized data matrix is expressed as

$$\mathbf{X} = \left(\mathbf{X} - \mathbf{1}\mu^\top\right)\Sigma^{-1}, \tag{3}$$

where $\mu \in \mathbb{R}^d$ collects the feature means, $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_d)$, and $\mathbf{1} \in \mathbb{R}^N$ denotes an all-ones vector. The normalized matrix $\mathbf{X}$ serves as the input to the LKRP module [13].

3.2.4. Temporal Segmentation

To enable sequential anomaly modeling, the normalized feature sequence is segmented into overlapping temporal windows. These windows are later used for both generative augmentation and sliding window anomaly detection, as described in the following Sections [12].

In aviation operations, variations in SNR, BER, and RSSI reflect changes in air–ground communication reliability. Abnormal temporal patterns in these indicators may correspond to reduced voice clarity, fading, or unstable sector transitions. The anomalies considered here therefore represent early signs of communication degradation.

3.3. Learnable Kernel–Redundancy Projection (LKRP)

To overcome the limitations of fixed-kernel projection and offline feature selection, the LKRP module learns a kernel-inspired nonlinear mapping that is optimized jointly with the downstream detection objective [13,15,16]. In aeronautical VHF monitoring, fixed kernel hyperparameters and offline selection may be insufficient under nonstationary channel conditions, extreme anomaly sparsity, and strict latency constraints [1–3].

Motivated by these limitations, the proposed Learnable Kernel–Redundancy Projection (LKRP) module adopts a kernel-inspired nonlinear projection that is jointly optimized with downstream detection objectives [13]. Unlike classical KPCA, LKRP learns

a compact projection that preserves salient structure while supporting discriminative anomaly detection.

3.3.1. Parametric Kernel Mapping

We introduce a parametric kernel operating on a learnable linear transformation of the input. Let $W \in \mathbb{R}^{d' \times d}$ be a trainable projection matrix and $\sigma > 0$ a trainable kernel bandwidth. The learnable RBF kernel is defined as

$$K_\theta(x_i, x_j) = \exp\left(-\frac{\|Wx_i - Wx_j\|_2^2}{\sigma^2}\right), \tag{4}$$

where $\theta = \{W, \sigma\}$ are learnable parameters. The kernel matrix becomes

$$K_\theta = [K_\theta(x_i, x_j)]_{i,j=1}^N \in \mathbb{R}^{N \times N}. \tag{5}$$

Kernel centering is performed as

$$\tilde{K}_\theta = K_\theta - \frac{1}{N}\mathbf{1}\mathbf{1}^\top K_\theta - \frac{1}{N}K_\theta\mathbf{1}\mathbf{1}^\top + \frac{1}{N^2}\mathbf{1}\mathbf{1}^\top K_\theta \mathbf{1}\mathbf{1}^\top. \tag{6}$$

Instead of eigen decomposition, projection directions are parameterized as a learnable matrix $U_\theta \in \mathbb{R}^{N \times m}$ optimized via gradient descent. The LKRP embedding is

$$Z = \tilde{K}_\theta U_\theta, \tag{7}$$

where each row $z_i \in \mathbb{R}^m$ represents the nonlinear embedding of frame x_i.

Efficient Inference via Anchor-Based Kernel Approximation

Direct kernel evaluation scales quadratically with the number of samples. To enable real-time deployment, LKRP adopts an anchor-based Nyström approximation (a low-rank kernel approximation method). During training, a fixed set of anchor points is learned from the training data. During inference, embeddings for streaming frames are computed only with respect to the anchors, reducing computational complexity from $O(N^2)$ to $O(NM)$ with $M \ll N$ anchors.

Implementation note: In practice, Z is computed via the centered kernel matrix without explicit reproducing kernel Hilbert space (RKHS) feature construction, enabling efficient end-to-end optimization.

3.3.2. Learnable Kernel Reconstruction Loss

To preserve nonlinear structure, a reconstruction-inspired regularizer is introduced. A linear decoder $V \in \mathbb{R}^{m \times d}$ maps embeddings back to the input space:

$$\hat{\mathbf{X}} = ZV^\top. \tag{8}$$

The learnable kernel loss is defined as

$$\mathcal{L}_{LK} = \|\mathbf{X} - \hat{\mathbf{X}}\|_F^2, \tag{9}$$

which encourages the kernel parameters and projection matrices to retain salient nonlinear structure in VHF features.

3.3.3. Differentiable MRMR Loss

Although the LKRP embedding is low-dimensional, some components may still be redundant or weakly informative. We therefore integrate a differentiable MRMR (DMR) objective [15,16]. Let $Z \in \mathbb{R}^{N\times m}$ and define a soft selection vector $w \in \mathbb{R}_+^m$ constrained on the simplex

$$\Delta_p = \{w \in \mathbb{R}_+^m : \mathbf{1}^\top w = p\}. \tag{10}$$

Mutual information is estimated using a neural estimator inspired by mutual information neural estimation (MINE) [37]. In practice, the joint distribution is obtained from paired samples within each mini-batch, while the product of marginals is approximated by randomly shuffling samples across the batch.

$$I_\psi(A;B) \approx \mathbb{E}_{p(a,b)}[T_\psi(a,b)] - \log \mathbb{E}_{p(a)p(b)}[\exp(T_\psi(a,b))]. \tag{11}$$

Using this estimator, relevance and redundancy are defined as

$$I_{fC} = [I_\psi(f_j;C)]_{j=1}^m, \qquad I_{ff} = [I_\psi(f_j;f_k)]_{j,k=1}^m. \tag{12}$$

The differentiable MRMR loss becomes

$$\mathcal{L}_{DMR} = -w^\top I_{fC} + \lambda_r w^\top I_{ff} w. \tag{13}$$

3.3.4. Optimization Strategy

Neural mutual-information estimation can be sensitive under extreme anomaly sparsity. To improve stability, T_ψ is implemented as a lightweight MLP and trained with mini-batch estimation and gradient clipping, while an exponential moving average stabilizes the log-partition term [37].

The selection weights are updated via projected gradient descent:

$$w^{(t+1)} = \Pi_{\Delta_p}\Big(w^{(t)} - \eta \nabla_w \mathcal{L}_{DMR}(w^{(t)})\Big), \tag{14}$$

where η is the learning rate and Π_{Δ_p} denotes projection onto the simplex Δ_p. After training, the top-p components according to w are retained to form the final LKRP feature vector used by subsequent modules.

The mutual-information relationships among the selected features, computed on the training set after LKRP pretraining, are visualized in Figure 3.

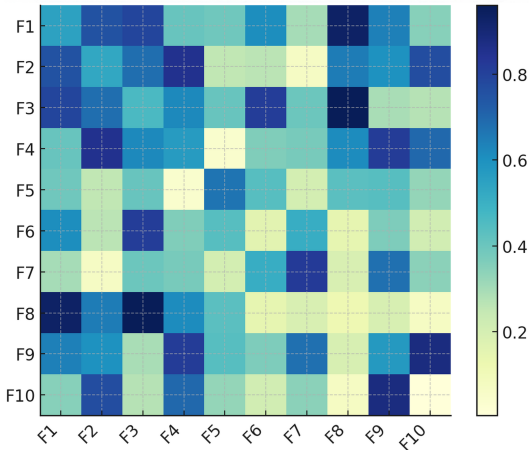


Figure 3. Mutual information matrix among the MRMR-selected features, computed on the training set after LKRP pretraining. The results show reduced redundancy and strong relevance of the retained LKRP components.

3.4. Conditional Temporal WGAN for Imbalance Mitigation

Severe class imbalance is a major challenge in aviation communication datasets. To increase anomaly coverage, a conditional temporal WGAN (CT–WGAN) is used to generate anomaly-like feature sequences conditioned on communication states [18,19,31,36]. The architecture of the proposed CT–WGAN is illustrated in Figure 4.

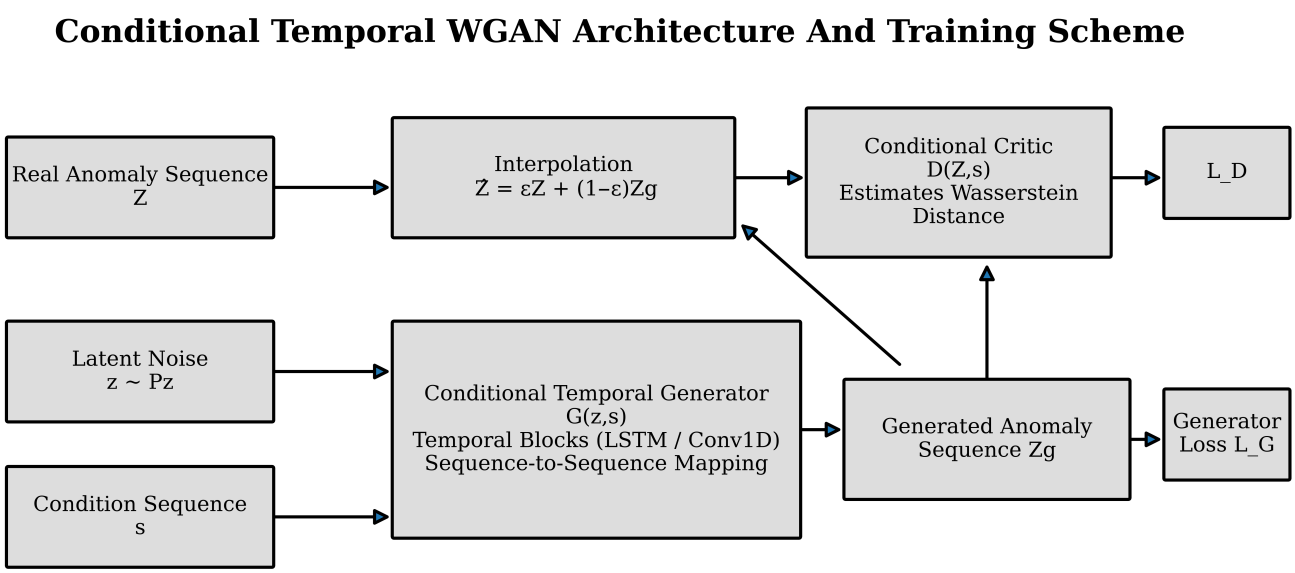


Figure 4. Architecture of the conditional temporal WGAN (CT–WGAN). The generator produces anomaly-like temporal sequences conditioned on communication-state variables, while the critic enforces temporal realism and condition consistency through Wasserstein adversarial training.

Let $\{Z_t\}_{t=1}^{T}$ denote a temporal sequence of LKRP features and $\{s_t\}_{t=1}^{T}$ the corresponding communication condition sequence (e.g., signal-to-noise ratio (SNR), received signal strength indicator (RSSI), bit error rate (BER), and packet error rate (PER)). The generator receives latent noise $z \sim P_z$ and condition sequence $s \sim P_s$ to synthesize anomaly-like sequences:

$$Z_g = G(z, s), \qquad z \sim P_z,\ s \sim P_s. \tag{15}$$

The critic takes both feature sequences and conditions as input, $D(Z, s)$, and the conditional Wasserstein distance is approximated as

$$W(P_r, P_g) \approx \sup_{\|D\|_L \leq 1} \mathbb{E}_{(Z,s)\sim P_r}[D(Z, s)] - \mathbb{E}_{(Z_g,s)\sim P_g}[D(Z_g, s)]. \tag{16}$$

The training dynamics and condition consistency of the proposed CT–WGAN are illustrated in Figure 5. The Wasserstein distance gradually decreases and stabilizes during training, indicating reliable adversarial convergence. The dynamic time warping (DTW)-based condition consistency analysis further shows that generated sequences follow diverse communication condition trajectories rather than repeating identical patterns.

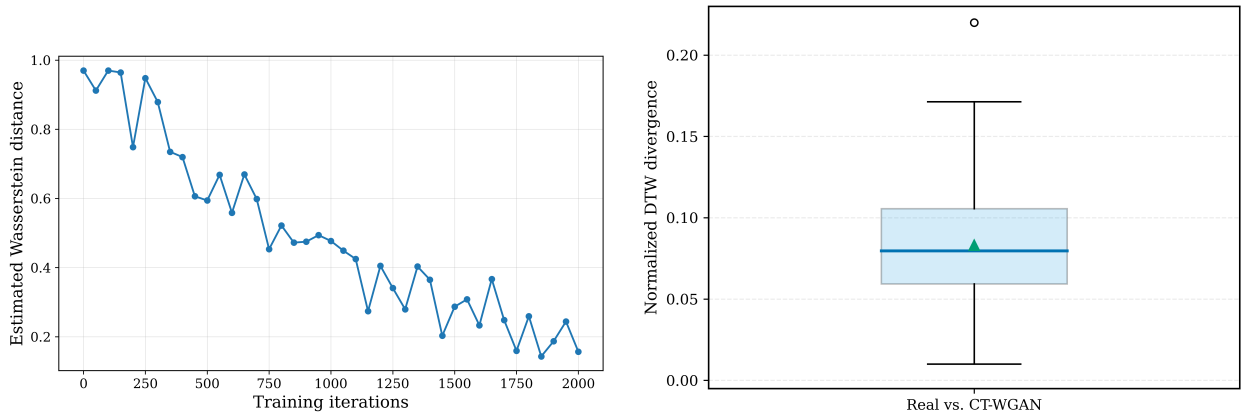


Figure 5. CT–WGAN training dynamics and condition consistency. (**Left**): Wasserstein distance convergence during training. (**Right**): DTW-based condition consistency between real and generated sequences. Circles denote real anomaly sequences and triangles denote generated sequences.

3.4.1. Gradient Penalty Regularization

To enforce the Lipschitz constraint and stabilize training, a gradient penalty is applied [19,31,36]. Interpolated samples are defined as

$$\hat{Z} = \epsilon Z + (1-\epsilon) Z_g, \qquad \epsilon \sim U(0,1). \tag{17}$$

The critic loss becomes

$$\mathcal{L}_D = -\mathbb{E}_{(Z,s)\sim P_r} D(Z,s) + \mathbb{E}_{(Z_g,s)\sim P_g} D(Z_g,s) + \lambda \mathbb{E}_{(\hat{Z},s)} \left(\|\nabla_{\hat{Z}} D(\hat{Z},s)\|_2 - 1 \right)^2. \tag{18}$$

The generator loss is

$$\mathcal{L}_G = -\mathbb{E}_{(Z_g,s)\sim P_g} D(Z_g,s). \tag{19}$$

Optimization Clarification

The critic and generator are optimized alternately by minimizing $\mathcal{L}_D$ and $\mathcal{L}_G$, respectively. Accordingly, $\mathcal{L}_{CT}$ in Equation (1) denotes the overall conditional adversarial training objective of this alternating optimization.

The conditional adversarial loss used in Equation (1) is therefore

$$\mathcal{L}_{CT} = \mathcal{L}_D + \mathcal{L}_G. \tag{20}$$

3.4.2. Training-Time Data Augmentation

After training, generated anomaly sequences are merged with real anomaly data to form an augmented training distribution:

$$P_{\text{aug}} = (1-\rho) P_r + \rho P_g, \tag{21}$$

where ρ controls the proportion of generated anomaly samples. Importantly, CT–WGAN is used exclusively during training; during inference the generative model is removed to ensure real-time deployment.

3.5. Hybrid CNN–LSTM Anomaly Detector

The final detection head adopts a lightweight hybrid CNN–LSTM architecture operating on LKRP+DMR-refined feature sequences. The CNN front-end captures multi-scale local temporal patterns, while the LSTM back-end models long-range dependencies and

evolving communication dynamics, providing a favorable accuracy–latency tradeoff for embedded VHF monitoring [12].

3.5.1. Sliding Window Construction

The construction of overlapping temporal windows used by the detector is illustrated in Figure 6.

Given the refined LKRP feature sequence $\{z_t\}_{t=1}^{T_{seq}}$ with $z_t \in \mathbb{R}^p$, overlapping windows of length T are constructed with stride s:

$$X_n = \{z_{t_n}, z_{t_n+1}, \ldots, z_{t_n+T-1}\} \in \mathbb{R}^{T\times p}, \qquad t_n = 1 + (n-1)s. \tag{22}$$

A window is labeled anomalous if it contains at least one anomalous frame. The window-level label is formally defined as

$$y_n = \max_{t\in\mathcal{W}_n} c_t, \tag{23}$$

where $\mathcal{W}_n$ denotes the set of frame indices contained in window X_n.

Window Labeling Rationale

This labeling rule follows the monitoring requirement that the presence of any anomalous frame within a window should raise an alert.

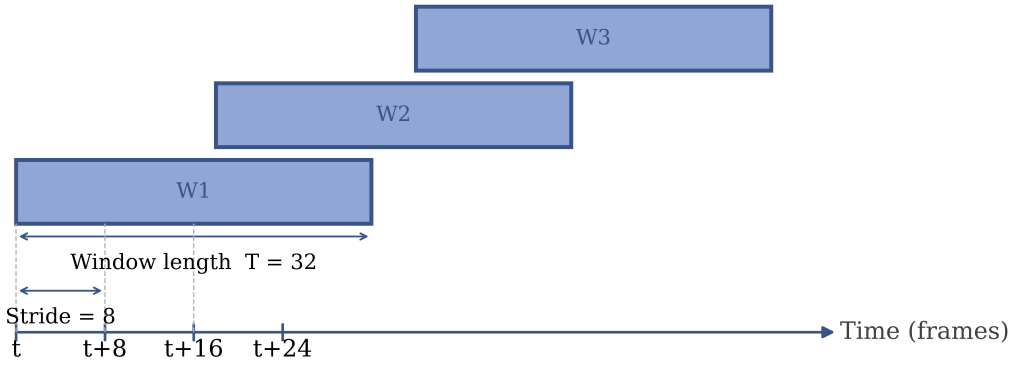


Figure 6. Sliding window construction for sequential anomaly detection. Overlapping windows are generated from LKRP feature sequences and labeled as anomalous if any frame within the window is abnormal.

3.5.2. Multi-Branch Temporal CNN

Each window is processed by parallel convolutional branches with kernel sizes $\{3, 5, 7\}$ to capture short-, medium-, and long-range temporal patterns:

$$h_{b,t} = \sigma\left(\sum_{\tau=0}^{k_b-1} W_{b,\tau} x_{t-\tau} + b_b\right). \tag{24}$$

Max pooling provides temporal invariance:

$$\tilde{h}_{b,j} = \max_{t\in\mathcal{N}_b(j)} h_{b,t}. \tag{25}$$

Outputs from all branches are concatenated:

$$H_{\text{CNN}} = [h^{(3)}; h^{(5)}; h^{(7)}] \in \mathbb{R}^{T'\times C}. \tag{26}$$

3.5.3. BiLSTM Classification

A bidirectional LSTM models long-range dependencies and outputs anomaly probabilities. Training uses class-weighted binary cross-entropy:

$$\mathcal{L}_{clf} = -\frac{1}{N_w}\sum_{n=1}^{N_w}\left[w_1 y_n \log \hat{y}_n + w_0(1-y_n)\log(1-\hat{y}_n)\right]. \tag{27}$$

3.5.4. Computational Complexity

The convolutional cost per window is approximated as

$$\text{FLOPs}_{\text{CNN}} \approx \sum_b T\, k_b\, c_{in}\, c_{out}. \tag{28}$$

The BiLSTM cost per window is approximated as

$$\text{FLOPs}_{\text{BiLSTM}} \approx 2T' \cdot 4h(C+h). \tag{29}$$

The fully connected layer cost per window is approximated as

$$\text{FLOPs}_{\text{FC}} \approx T'C. \tag{30}$$

The total detector complexity is

$$\text{FLOPs}_{\text{det}} = \text{FLOPs}_{\text{CNN}} + \text{FLOPs}_{\text{BiLSTM}} + \text{FLOPs}_{\text{FC}}. \tag{31}$$

Only LKRP and CNN–LSTM are used during inference, ensuring real-time deployment feasibility.

3.6. Summary of the Proposed Framework

The proposed method combines three complementary components for VHF communication monitoring. The LKRP module learns a nonlinear feature space while reducing redundancy in the learned representations. The CT–WGAN expands minority anomaly samples during training by generating condition-consistent temporal sequences. The CNN–LSTM detector operates on sliding window LKRP features to produce real-time anomaly predictions.

During deployment, only LKRP and the CNN–LSTM detector are executed, ensuring low computational cost and predictable runtime behavior. This design allows the model to handle nonlinear signal characteristics, temporal dependencies, and extreme class imbalance while remaining suitable for real-time avionics monitoring.

4. Experiments

This section describes how the proposed LKRP–CT–WGAN framework is evaluated. We introduce the dataset, baseline methods, evaluation metrics, implementation details, and the overall experimental procedure. The experimental design follows commonly adopted practices in PHM-oriented anomaly detection under severe class imbalance [5,6] and follows experimental settings commonly used in communication and deep anomaly-detection studies [4,7,8,12,38].

4.1. Dataset Description

The dataset contains approximately 1.2×10^6 VHF communication frames, of which fewer than 1% are labelled as anomalous. This level of imbalance is common in operational monitoring data [3,5,7]. The evaluation therefore focuses on imbalance-aware metrics

such as F1-score, PR-AUC, and ROC-AUC rather than accuracy alone [7,8]. The data were used exclusively for monitoring research and do not include safety-critical control decisions. All data were anonymized and processed in accordance with aviation data-protection requirements.

4.1.1. Real-World VHF Dataset

The operational data come from civil flights spanning roughly 2000 flight hours. The recordings include VHF communication logs together with ground station monitoring data. Each frame includes link quality indicators such as SNR, BER, RSSI, PER, retransmission count, and transmission latency, which are commonly used in wireless anomaly-detection studies [2,3].

4.1.2. Simulated Degradation Anomalies

Because real anomalies are limited, additional degradation scenarios were generated using a channel emulator. These scenarios include fading, multipath interference, Doppler shifts [1], amplifier degradation, and additive noise. The use of both operational and simulated anomalies helps broaden the coverage of representative degradation patterns during training when real samples are scarce [1–3].

4.1.3. Anomaly Definition and Labeling

A frame is marked as anomalous when abnormal communication-quality patterns appear, such as sustained BER/PER spikes, abrupt RSSI drops, or bursts in latency and retransmissions that exceed operational thresholds recorded in the monitoring logs. For simulated data, the anomaly labels are inherited directly from the injected degradation scenarios.

Figure 7 shows representative temporal trajectories of key communication indicators and illustrates the coupled degradation behavior that motivates the use of temporal modeling.

4.1.4. Leakage-Aware Data Split

To avoid temporal leakage, the dataset is partitioned into training (60%), validation (20%), and test (20%) splits at the flight/session level (i.e., all frames from the same flight/session belong to exactly one split) [7,8]. Anomalies are present in all splits, but no anomaly segments are shared across splits.

4.1.5. Dataset Statistics

Table 1 summarizes key dataset statistics and the sliding window parameters used throughout the experiments.

Table 1. Summary of dataset and sliding window statistics.

Item	Value
Total number of frames	$\approx 1.2 \times 10^6$
Number of features per frame (d)	6
Anomalous frames ratio	$<1\%$
Flight/session duration	≈2000 flight hours
Train/Validation/Test split	60%/20%/20%
Sliding window length (T)	32 frames
Sliding window stride (s)	8 frames
LKRP embedding dimension (m)	50
DMR retained components (p)	20

The monitoring logs were recorded at an effective frame rate of approximately 5–10 Hz after preprocessing and quality filtering.

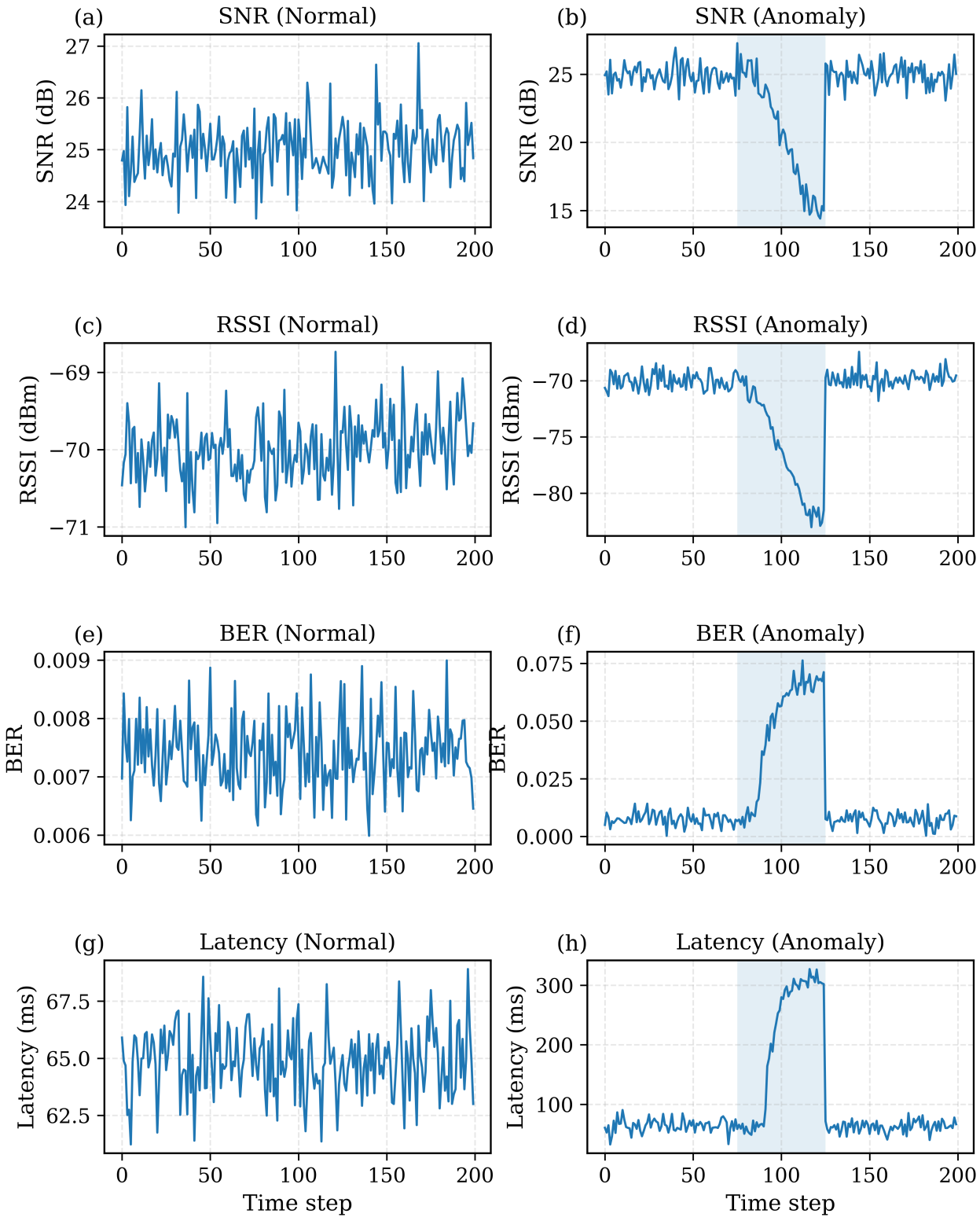


Figure 7. (**a**,**b**) Signal-to-noise ratio (SNR), (**c**,**d**) received signal strength indicator (RSSI), (**e**,**f**) bit error rate (BER), and (**g**,**h**) end-to-end latency. The left column corresponds to normal communication, and the right column shows an anomaly interval (shaded region). During the anomaly, RSSI decreases, followed by a delayed reduction in SNR, a rapid increase in BER, and a pronounced rise in latency.

4.1.6. Proportion of Operational vs. Simulated Anomalies

In civil aviation monitoring, naturally occurring communication anomalies are extremely rare and cannot provide sufficient coverage for training deep models. To improve anomaly diversity during training, simulated degradation scenarios are therefore incorporated.

Specifically, anomaly samples in the training set consist of approximately 80% simulated degradation anomalies and 20% operational anomalies. Simulated anomalies are used only during training, while the validation and test sets contain operational (real-world) anomalies exclusively.

This protocol evaluates generalization from mixed training anomalies to purely real anomalies. Although simulated degradations cannot reproduce all real-world conditions, they provide controlled coverage of representative failure mechanisms and help improve robustness when real anomaly samples are limited.

The composition of operational and simulated anomalies across dataset splits is summarized in Table 2.

Table 2. Composition of operational vs. simulated anomalies.

Split	Operational Anomalies	Simulated Anomalies	Notes
Training	≈20%	≈80%	used for model training
Validation	100%	0%	real anomalies only
Test	100%	0%	real anomalies only

Besides domain-specific aviation data, public RF datasets such as CommRad RF provide useful reference benchmarks for communication signal detection and classification studies [39].

4.2. Baseline Methods

We compare the proposed approach with representative classical and deep learning baselines that are commonly used in communication anomaly detection and PHM-oriented learning [5,7,8]:

- SVM (RBF) trained on standardized raw features.
- Random Forest (RF) with 200 trees and Gini splitting.
- PCA–LR: linear PCA on raw features followed by logistic regression.
- Autoencoder (AE) trained to reconstruct raw frames, using reconstruction error for anomaly scoring.
- LSTM Autoencoder (LSTM–AE) for sequence reconstruction-based anomaly detection.
- Variational Autoencoder (VAE) as a probabilistic generative anomaly detector.

4.2.1. Transformer-Based Baseline

To compare with recent deep sequential anomaly detection models, we additionally implement a lightweight Temporal Transformer baseline, following established transformer-based anomaly detection practices in time series and wireless settings [40]. The model uses the same sliding window inputs as the CNN–LSTM detector for fair comparison [41]. The sliding-window segmentation used in this work is illustrated in Figure 6.

4.2.2. Ablation Variants

To quantify the contribution of each proposed module, we report:

- CNN–LSTM (Raw): trained on normalized raw features without LKRP, DMR, or CT–WGAN.
- LKRP(+DMR) + CNN–LSTM: trained on LKRP(+DMR) features without CT–WGAN augmentation.

- Full model (LKRP + CT–WGAN + CNN–LSTM): the complete framework.

4.3. Evaluation Metrics

Given the extreme imbalance, we report metrics beyond accuracy [7,8]. Let TP, TN, FP, and FN denote confusion matrix entries:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{32}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{33}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{34}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{35}$$

4.3.1. ROC-AUC and PR-AUC

We compute ROC-AUC by varying the decision threshold on the predicted anomaly probability. Because ROC curves can be optimistic under strong imbalance, we also report the area under the Precision–Recall curve (PR-AUC), which is more informative for rare event detection [7,8].

4.3.2. False Alarm Rate (FAR) and Low-FAR Operating Points

Safety-critical monitoring is sensitive to false alarms. We report

$$\text{FAR} = \frac{FP}{FP + TN}, \tag{36}$$

and additionally evaluate detection performance under strict false-alarm constraints by reporting the true positive rate at a fixed false alarm rate (FAR = 1%), denoted as $\text{TPR} \,|\, \text{FAR} = 1\%$. This operating point reflects practical monitoring settings where excessive false alarms may increase controller workload and reduce trust in automated monitoring tools.

4.3.3. Latency

Detection latency is measured as average inference time per test frame:

$$\text{Latency} = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} t_i, \tag{37}$$

where t_i is the processing time for the i-th test frame.

4.4. Implementation Details

Implementation follows established practices for learnable kernel representations [13], neural mutual-information estimation [37], adversarial augmentation [19,31,36], and hybrid deep temporal modeling [12].

4.4.1. LKRP Configuration

The parametric RBF kernel in (4) initializes the bandwidth σ using a median heuristic and updates it during training. The embedding dimension is set to $m = 50$. The reconstruction-inspired loss $\mathcal{L}_{LK}$ (cf. Equation (9)) is weighted by β.

4.4.2. Differentiable MRMR (DMR)

The selection vector w is initialized uniformly on the simplex with $p = 20$ effective components retained. Mutual information terms are estimated using a lightweight neural estimator (three-layer MLP) as in MINE [37], with mini-batch estimation and gradient clipping for stability. The DMR loss $\mathcal{L}_{DMR}$ is weighted by γ and is motivated by MRMR-style relevance–redundancy principles [15,16].

4.4.3. CT–WGAN

We adopt WGAN-GP training with gradient penalty $\lambda = 10$ and five critic updates per generator step for stability [19,31,36]. The model is trained for 50 epochs. CT–WGAN is used only during training to generate condition-consistent anomaly sequences, following conditional augmentation practice in wireless anomaly detection [2,3].

4.4.4. CNN–LSTM Detector and Windowing

The detector uses a multi-branch 1D CNN front-end with kernel sizes $\{3, 5, 7\}$ and max pooling, followed by a BiLSTM with 128 hidden units per direction. Sliding window parameters are $T = 32$ and $s = 8$.

4.4.5. Optimization and Training Schedule

All neural components are trained using Adam ($\alpha = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$) with batch size 128. We adopt a two-stage strategy: (i) pretrain LKRP+DMR using $\mathcal{L}_{LK} + \gamma\mathcal{L}_{DMR}$ and train CT–WGAN using $\mathcal{L}_{CT}$; (ii) train the CNN–LSTM detector on the augmented dataset and then jointly fine-tune LKRP and the detector under the full loss in Equation (1). Early stopping is applied based on validation F1-score.

4.4.6. Hardware Platform

Training experiments were conducted on an NVIDIA RTX 3090 GPU. Inference latency was measured in a CPU-only setting to reflect deployment in ground station monitoring systems. All latency measurements correspond to single-sample inference (batch size = 1) on an Intel Xeon-class CPU.

4.5. Experimental Procedure

The overall workflow follows the LKRP–CT–WGAN–CNN–LSTM pipeline (Figure 2):

1. Preprocessing and splitting: Normalize frame-level features and partition the dataset using leakage-aware flight/session-level splitting.
2. Representation learning (LKRP+DMR): Train LKRP with $\mathcal{L}_{LK}$ and $\mathcal{L}_{DMR}$ to obtain compact, redundancy-suppressed embeddings.
3. Conditional temporal augmentation (CT–WGAN): Train CT–WGAN on anomaly sequences in LKRP space and their condition sequences to generate condition-consistent synthetic anomalies [19,31,36].
4. Detector training: Construct sliding windows over LKRP+DMR feature sequences and augment the minority class with CT–WGAN-generated anomaly windows. Train the CNN–LSTM detector with $\mathcal{L}_{clf}$.
5. Fine-tuning and evaluation: Jointly fine-tune LKRP and the detector under Equation (1) and evaluate all baselines and variants on the held-out test set using the metrics in Section 4.3.

4.5.1. Threshold Selection

For all probabilistic detectors, the decision threshold is selected on the validation set to maximize F1-score and then applied to the held-out test set.

4.5.2. Reproducibility

All experiments are repeated $R = 5$ times with different random seeds. For a metric M, we report mean and standard deviation:

$$\overline{M} = \frac{1}{R}\sum_{r=1}^{R} M^{(r)}, \tag{38}$$

$$\mathrm{Std}(M) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(M^{(r)} - \overline{M}\right)^2}. \tag{39}$$

4.6. Reproducibility and Data Availability

The operational VHF communication dataset used in this study contains safety-critical aviation data collected from routine flight operations and ground monitoring systems. Due to aviation safety regulations and confidentiality constraints, the raw dataset cannot be publicly released.

To facilitate reproducibility and future research, we will publicly release the following resources upon publication:

- The complete model implementation of the LKRP–CT–WGAN–CNN–LSTM framework.
- All training and evaluation scripts.
- Hyperparameter configurations and preprocessing procedures.
- A synthetic benchmark dataset that preserves the statistical characteristics of the operational data.
- Detailed instructions to reproduce all experiments and figures reported in this paper.

These materials will enable independent verification of the proposed methodology while respecting aviation data security requirements.

5. Results and Discussion

This section presents the main experimental results of the proposed LKRP + CT–WGAN + CNN–LSTM framework. We first compare overall detection performance against classical and deep baselines, and then examine the impact of individual components through ablation results. All reported results are summarized as mean ± standard deviation over five runs with different random seeds, which is a common evaluation practice in anomaly detection studies [7] and PHM-oriented research [8].

5.1. Overall Detection Effectiveness

Table 3 reports the performance of all benchmark methods. The proposed framework achieves the best average results, with an F1-score of 0.947 ± 0.003 and an AUC of 0.972 ± 0.002. Relative to the strongest baseline (VAE), the proposed method increases F1 by 0.107 (absolute) and AUC by 0.061, indicating improved separability under rare event conditions. The model also attains precision 0.932 and recall 0.963, which reduces missed anomalies while keeping false alarms under control—a key requirement in safety-critical aviation monitoring [3,7]. Overall, the gains are consistent across classical machine learning baselines (SVM, RF, PCA–LR) and deep learning baselines (AE, LSTM–AE, VAE). These results indicate that learnable nonlinear representations together with redundancy-aware feature shaping improve class separability [13,15,16]. The hybrid temporal detector also benefits from the resulting structured feature space [12]. In addition, the use of condition-consistent augmentation helps improve detection performance under severe class imbalance [18,19,31,36].

Table 3. Overall performance comparison of anomaly detection methods (mean ± std over 5 runs). Bold values indicate the best performance across all methods.

Method	Acc	Prec.	Rec.	F1	AUC
SVM (RBF)	0.872 ± 0.006	0.801 ± 0.008	0.734 ± 0.010	0.766 ± 0.009	0.852 ± 0.007
Random Forest	0.889 ± 0.005	0.824 ± 0.007	0.761 ± 0.009	0.791 ± 0.008	0.861 ± 0.006
PCA–LR	0.864 ± 0.007	0.777 ± 0.009	0.708 ± 0.011	0.741 ± 0.010	0.835 ± 0.008
Autoencoder	0.903 ± 0.004	0.842 ± 0.006	0.781 ± 0.008	0.810 ± 0.007	0.884 ± 0.005
LSTM–AE	0.917 ± 0.004	0.859 ± 0.006	0.803 ± 0.007	0.830 ± 0.006	0.896 ± 0.005
VAE	0.924 ± 0.003	0.867 ± 0.005	0.814 ± 0.006	0.840 ± 0.005	0.911 ± 0.004
Proposed	**0.956 ± 0.002**	**0.932 ± 0.004**	**0.963 ± 0.003**	**0.947 ± 0.003**	**0.972 ± 0.002**

A visual comparison of F1-score and AUC is provided in Figure 8.

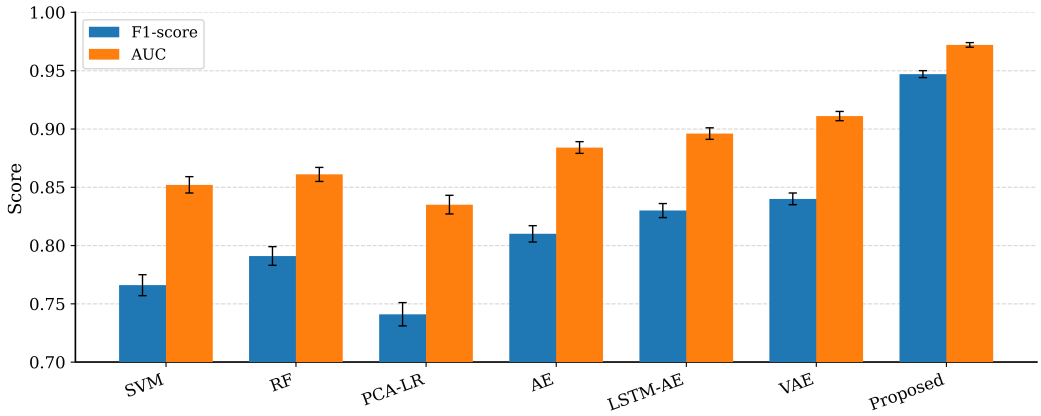


Figure 8. Overall comparison of anomaly detection methods in terms of F1-score and AUC.

5.2. Effect of Imbalance Mitigation Strategies

The VHF monitoring dataset used in this study is highly imbalanced, with anomalous frames accounting for less than 1% of the observations. Although CT–WGAN is designed to address this issue, it remains necessary to verify whether the observed performance gain can be explained by simpler imbalance handling techniques.

To clarify this point, several commonly used strategies are evaluated while keeping the CNN–LSTM detector unchanged. This experiment isolates the effect of the imbalance handling method from the detection architecture.

The following training strategies are considered:

1. No augmentation: the detector is trained using the original imbalanced dataset.
2. Class-weighted BCE: anomaly samples are assigned higher loss weights.
3. Synthetic Minority Over-sampling Technique (SMOTE) oversampling: minority samples are interpolated in the feature space.
4. WGAN-GP augmentation: anomaly samples are generated using an unconditional WGAN-GP model [18,19].
5. Proposed CT–WGAN: anomaly sequences are generated using the proposed conditional temporal model.

Table 4 summarizes the results. Applying a class-weighted loss increases recall, but this improvement comes with a higher number of false alarms, leading to reduced precision. SMOTE provides only limited gains, suggesting that simple interpolation does not adequately capture the temporal structure of communication anomalies.

Table 4. Comparison of imbalance mitigation strategies using the same CNN–LSTM detector (mean ± std over 5 runs). Bold values indicate the best performance among all strategies.

Training Strategy	Precision	Recall	F1	AUC
No augmentation	0.914 ± 0.006	0.823 ± 0.008	0.866 ± 0.007	0.941 ± 0.005
Class-weighted BCE	0.889 ± 0.007	0.891 ± 0.006	0.890 ± 0.006	0.951 ± 0.004
SMOTE oversampling	0.902 ± 0.006	0.872 ± 0.007	0.887 ± 0.006	0.953 ± 0.004
WGAN-GP augmentation	0.921 ± 0.005	0.912 ± 0.006	0.916 ± 0.005	0.962 ± 0.003
Proposed CT–WGAN	$\mathbf{0.932 \pm 0.004}$	$\mathbf{0.963 \pm 0.003}$	$\mathbf{0.947 \pm 0.003}$	$\mathbf{0.972 \pm 0.002}$

In contrast, WGAN-GP augmentation produces a more balanced improvement across metrics, indicating that generative augmentation is more suitable for this task. The proposed CT–WGAN achieves the strongest overall performance. The improvement is largely associated with temporal sequence generation and the use of communication-state conditioning, which leads to more realistic anomaly sequences during training.

5.3. Generalization to Unseen Anomaly Scenarios

In operational aviation environments, previously unseen communication degradation patterns may appear. It is therefore important to assess whether the detector can identify anomaly types that are not present during training.

To evaluate this capability, a leave-one-scenario-out experiment is performed. The dataset includes multiple anomaly sources, including multipath interference, electromagnetic interference, and hardware degradation. For each run, one anomaly category is excluded from training and used only for testing, creating an unseen anomaly evaluation setting.

Table 5 shows that detection performance remains high even when the anomaly type is not seen during training. The F1-score is slightly lower than in the standard test setting, but the average remains above 0.92, showing consistent performance across unseen anomaly categories.

Table 5. Generalization to unseen anomaly scenarios (F1-score). Bold values indicate the overall average performance across unseen anomaly types.

Unseen Anomaly Type	F1-Score
Multipath interference	0.931 ± 0.004
Electromagnetic interference	0.924 ± 0.005
Hardware degradation	0.918 ± 0.006
Average	$\mathbf{0.924 \pm 0.005}$

5.4. Robustness to Noisy Condition Inputs

In real deployments, communication-state variables such as SNR, RSSI, and BER may contain measurement noise or missing entries.

We evaluate robustness by introducing controlled perturbations to the condition vectors during testing. Two perturbation types are considered: (i) Additive Gaussian noise with increasing variance; (ii) Randomly missing condition entries replaced by their mean values.

As shown in Table 6, performance decreases gradually as the condition vectors become noisier or partially missing. Even with 40% of the condition values unavailable, the F1-score remains above 0.92.

Table 6. Robustness to noisy condition vectors.

Condition Perturbation	F1-Score
No noise	0.947 ± 0.003
Gaussian noise ($\sigma = 0.05$)	0.941 ± 0.004
Gaussian noise ($\sigma = 0.10$)	0.936 ± 0.004
Missing 20% conditions	0.933 ± 0.005
Missing 40% conditions	0.927 ± 0.006

5.5. Ablation Study

Table 7 and Figure 9 report the ablation results for individual components. Removing CT–WGAN leads to the largest performance drop (F1: $0.947 \rightarrow 0.897$; AUC: $0.972 \rightarrow 0.921$), showing the impact of temporal data augmentation when anomalies are extremely rare [19,31,36]. Disabling LKRP or DMR also reduces performance (Without LKRP: F1 = 0.911; Without DMR: F1 = 0.923), indicating that adaptive nonlinear projection and redundancy suppression both improve separability under nonstationary VHF conditions [13,15,16,37]. Training the CNN–LSTM detector directly on raw features yields the lowest performance (F1 = 0.874), showing the importance of representation shaping before sequential modeling.

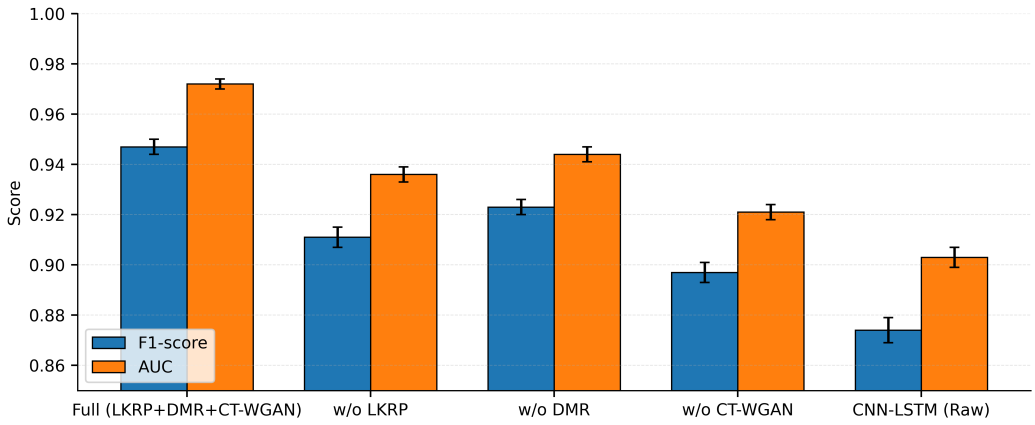


Figure 9. Ablation study results measured by F1-score and AUC. The full model shows the highest and most stable performance. Removing LKRP, DMR, or CT–WGAN leads to noticeable degradation. Error bars denote standard deviation.

Table 7. Ablation study results (mean ± standard deviation over five runs).

Configuration	F1	AUC
Full Model (LKRP + DMR + CT–WGAN)	0.947 ± 0.003	0.972 ± 0.002
Without CT–WGAN	0.897 ± 0.004	0.921 ± 0.003
Without DMR	0.923 ± 0.003	0.944 ± 0.003
Without LKRP	0.911 ± 0.004	0.936 ± 0.003
CNN–LSTM Raw	0.874 ± 0.005	0.903 ± 0.004

5.6. Detector Training Convergence

Figure 10 shows the training and validation loss curves of the CNN–LSTM detector. Both curves decrease during training and gradually stabilize in the later epochs. The gap between them remains small, suggesting limited overfitting. Early stopping is selected according to the validation F1-score, and the final model is taken from the best validation epoch.

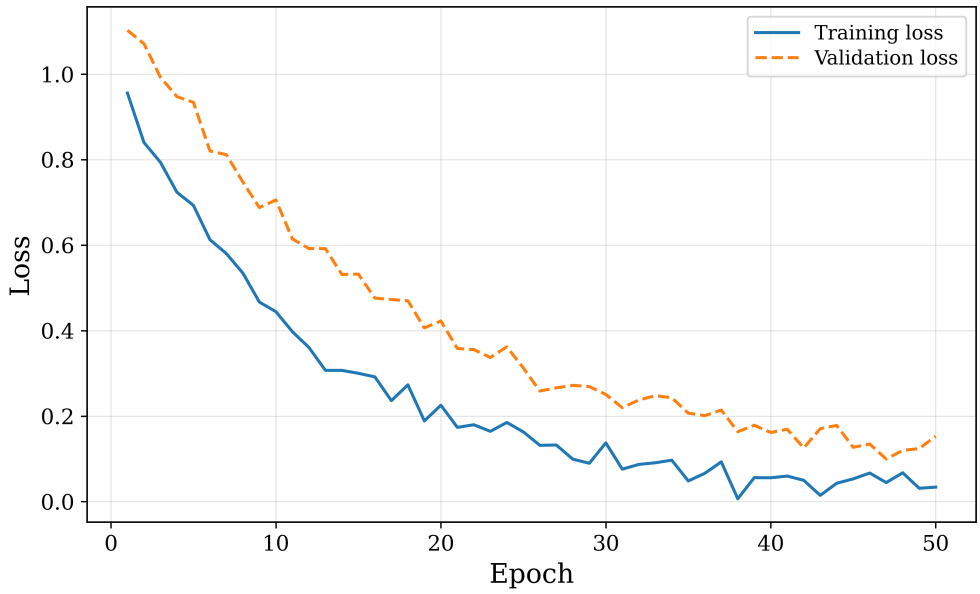


Figure 10. Training and validation loss curves for the CNN–LSTM detector.

5.7. Interpretability of Detected Anomalies

Temporal anomaly scores from the CNN–LSTM model are examined on several representative flight segments. Elevated scores are observed during periods of rapid SNR variation, sudden RSSI drops, bursty BER behavior, and short-term signal fading. These patterns are consistent with known VHF communication degradation phenomena and indicate sensitivity to temporal dynamics rather than isolated frame-level outliers.

5.8. ROC, PR, and Confusion Matrix Analysis

Confusion matrix results are evaluated on the window-level test set, where each sample corresponds to a sliding window segment. To avoid temporal leakage, the data are split at the flight/session level, and only windows from unseen flights are used for the final evaluation. The test set contains approximately 460 windows, of which about 30% are anomalous. Because a window is labeled anomalous if it contains at least one anomalous frame, the anomaly ratio increases at the window level even though frame-level anomalies are below 1%.

To assess performance under severe class imbalance, we examine ROC curves, Precision–Recall curves, and the confusion matrix [3,7,8]. The ROC curve characterizes the trade-off between the true positive rate and the false positive rate, while the PR curve is more informative for rare event detection.

Figure 11 shows that the proposed framework achieves consistently higher true positive rates across the full FPR range and obtains the largest ROC-AUC. Figure 12 shows that the method maintains higher precision over the recall range and achieves the best PR-AUC. The confusion matrix in Figure 13 shows limited false negatives while false positives remain controlled, supporting suitability for real-time deployment [7,8].

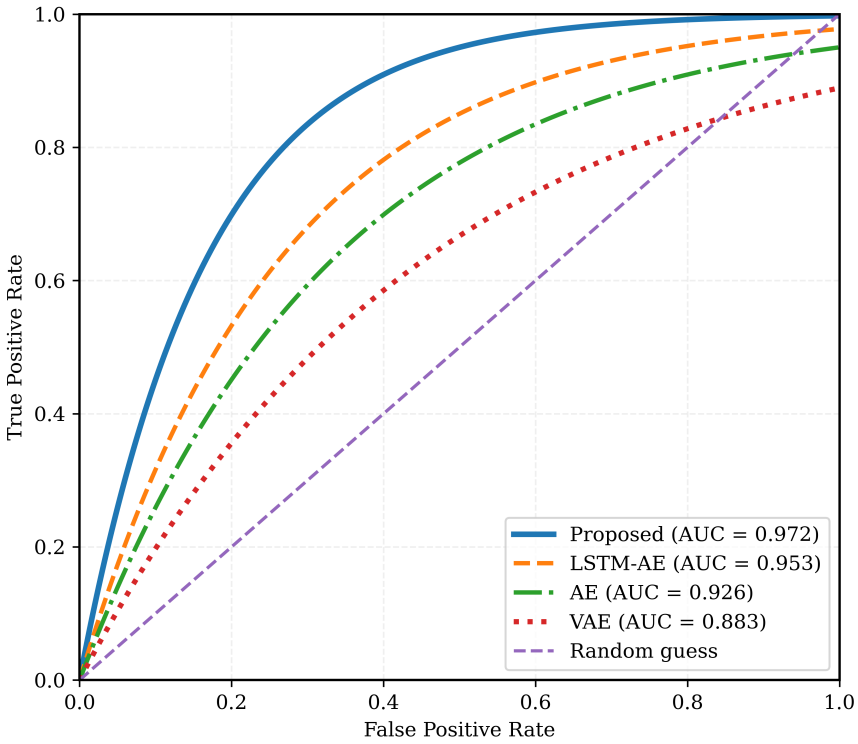


Figure 11. ROC curves comparing the proposed framework with baseline methods. The proposed model achieves the highest ROC-AUC. The diagonal dashed line corresponds to a random classifier.

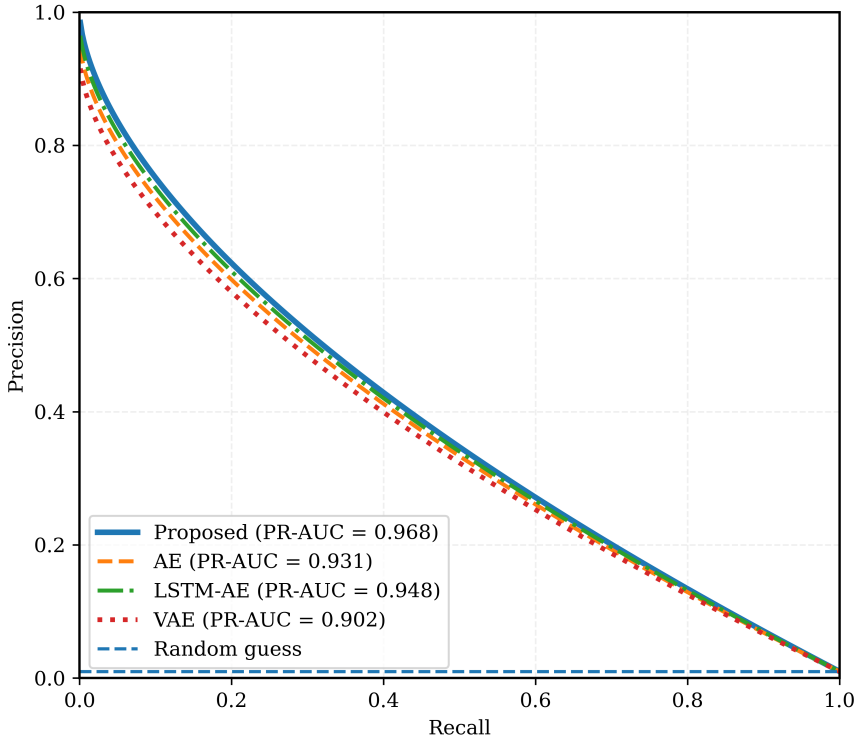


Figure 12. Precision–Recall curves comparing the proposed framework with baseline methods. The proposed model achieves the highest PR-AUC and maintains superior precision across the full recall range. The dashed horizontal line denotes the performance of a random classifier under the observed anomaly ratio.

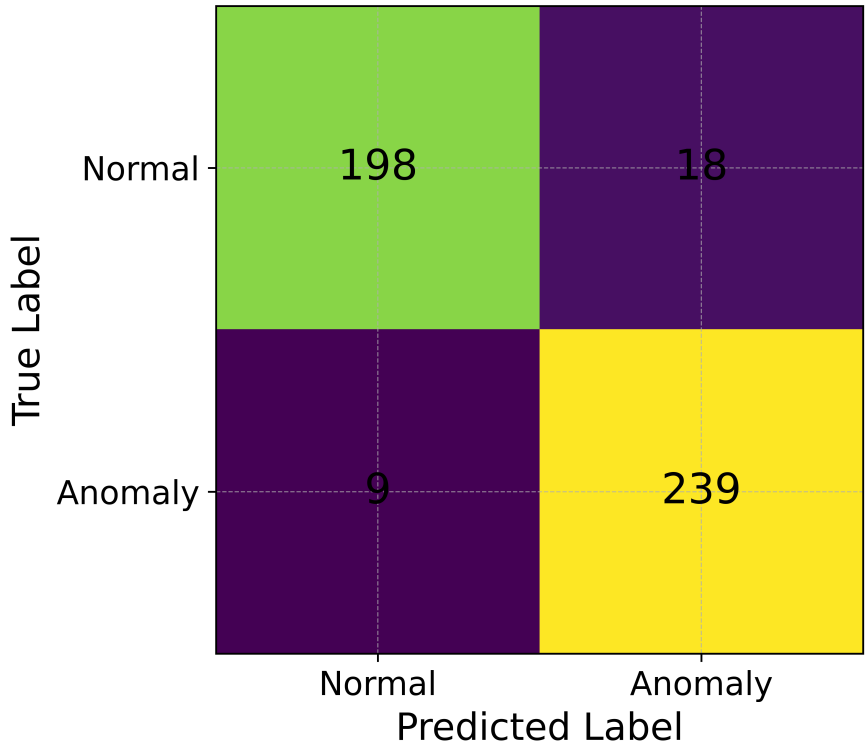


Figure 13. Confusion matrix of the proposed framework on the window-level test set.

5.9. Understanding the Proposed Framework

To provide additional insight into the behavior of the proposed method, we examine the generative augmentation module, the sensitivity to key hyperparameters, and the role of redundancy suppression in the learned representation.

5.9.1. CT–WGAN Generation Quality

Figure 5 shows the training behavior of CT–WGAN and the consistency of generated sequences under different communication conditions. The Wasserstein distance decreases and stabilizes during training, indicating stable adversarial optimization. The DTW-based comparison further shows that generated sequences follow the conditioning variables and retain meaningful temporal dynamics [19,31,36].

To assess whether the generated sequences reproduce the statistical properties of real anomalies, Figure 14 compares real and generated sequences from several perspectives, including time-domain behavior, marginal distributions, mean temporal profiles, autocorrelation functions, and frequency-domain energy spectra. The generated anomalies follow the overall patterns of real anomalies while retaining small variations, indicating that CT–WGAN captures key temporal and spectral characteristics without simply replicating the training data.

Figure 15 offers further insight into the generated anomaly samples. The ACF comparison shows that temporal correlations are retained, while the diversity–fidelity analysis reveals that the generated sequences span a wide range of patterns instead of clustering in a narrow region of the feature space. This indicates that CT–WGAN produces samples that are both realistic and varied, which helps enlarge minority-class coverage and supports detection under severe class imbalance [19,31,36].

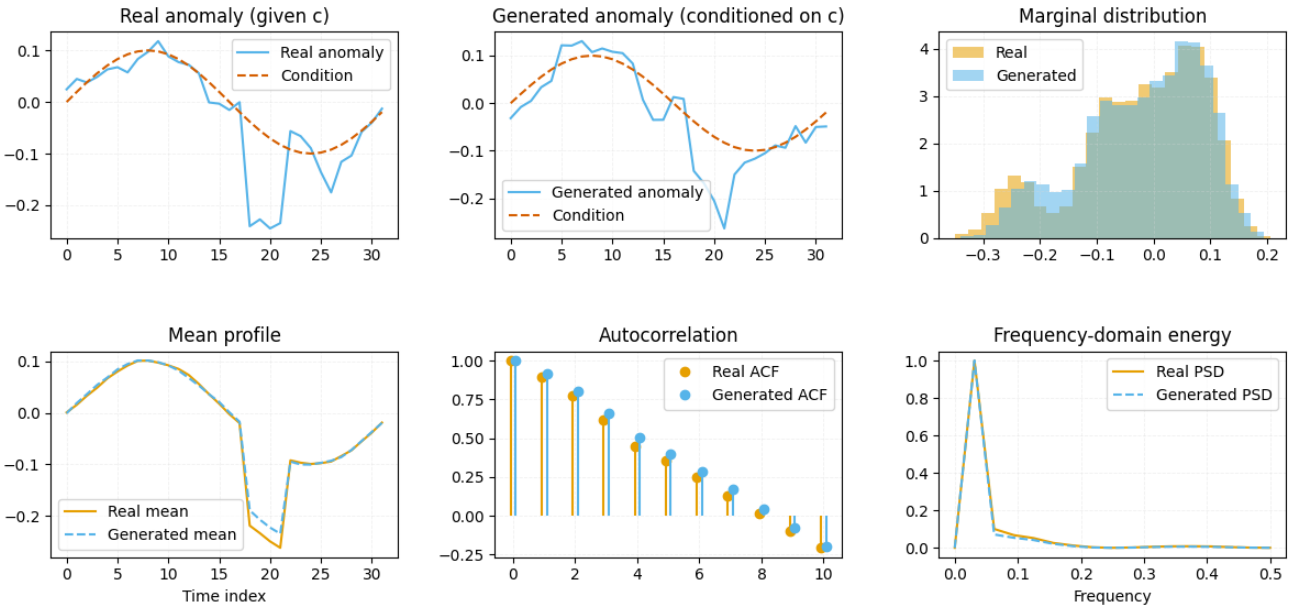


Figure 14. Statistical comparison between real and CT–WGAN generated anomaly sequences. Real and generated samples are compared in the time domain, distributional statistics, temporal correlations, and frequency-domain energy. The generated anomalies closely follow the real anomaly characteristics while maintaining slight variations, indicating realistic yet non-identical generation.

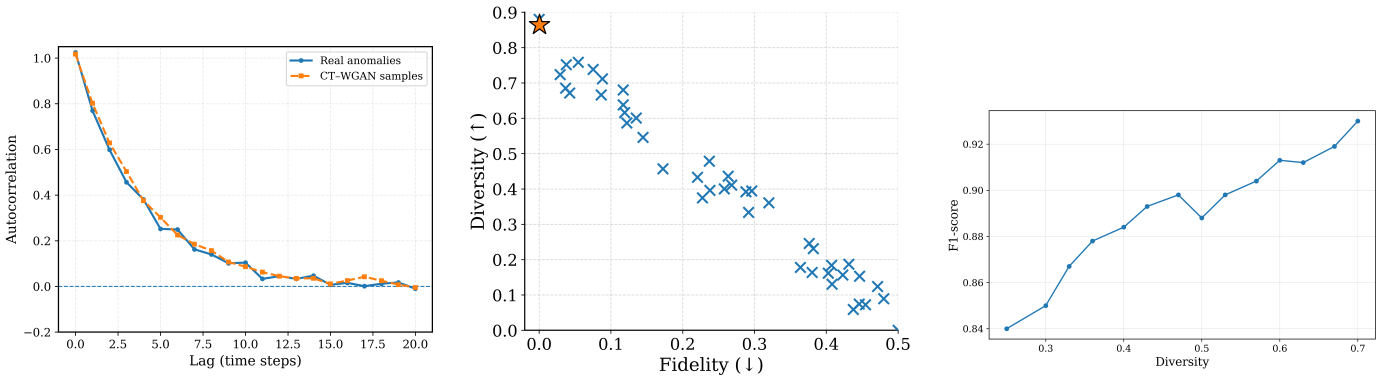


Figure 15. Quality assessment of CT–WGAN generated anomalies: ACF consistency (**left**), diversity–fidelity scatter (**middle**), and Diversity–F1 Pareto frontier (**right**). The dotted line in the left panel indicates the reference autocorrelation of real sequences. The asterisk in the middle panel denotes the selected operating point, and the arrow indicates the Pareto improvement direction.

Mode Collapse Analysis

Mode collapse is a common concern in GAN-based augmentation, where the generator may reproduce only a limited subset of patterns. Although direct verification is challenging for real-world communication data, several observations indicate that severe memorization is unlikely. The diversity–fidelity distribution in Figure 15 shows a broad spread of generated samples rather than tight clustering. In addition, the performance gains observed in the ablation study indicate that the synthetic data expand anomaly coverage instead of duplicating existing samples. A complete characterization of generative diversity for rare aviation anomalies remains an open research problem.

5.9.2. Hyperparameter Sensitivity and Interaction

To evaluate robustness to key hyperparameters, we analyze the sensitivity of the model to the LKRP kernel width σ, the CT–WGAN gradient penalty coefficient λ, and the sliding window length T. The sensitivity curves are shown in Figure 16. In addition,

the joint effects of key parameter pairs are visualized in Figure 17. Overall, the model maintains stable performance across a wide parameter range, indicating that the proposed framework does not rely on fragile tuning.

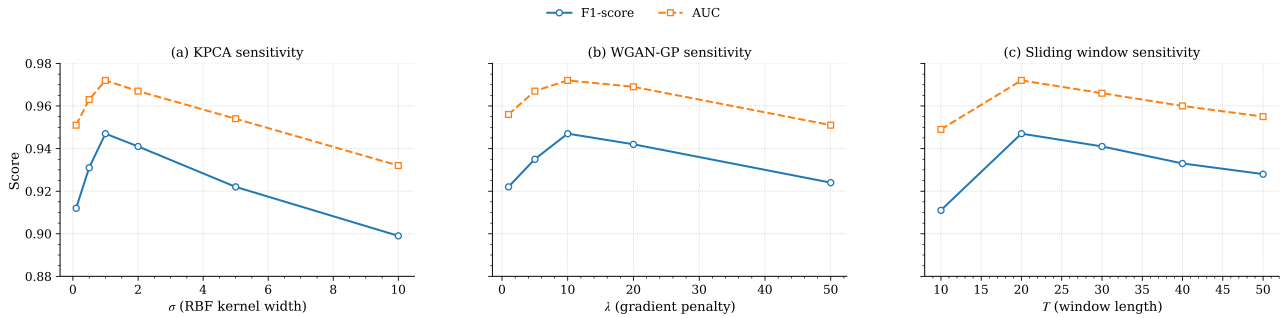


Figure 16. Sensitivity analysis of key hyperparameters, including LKRP kernel width σ, CT–WGAN penalty coefficient λ, and sliding window length T.

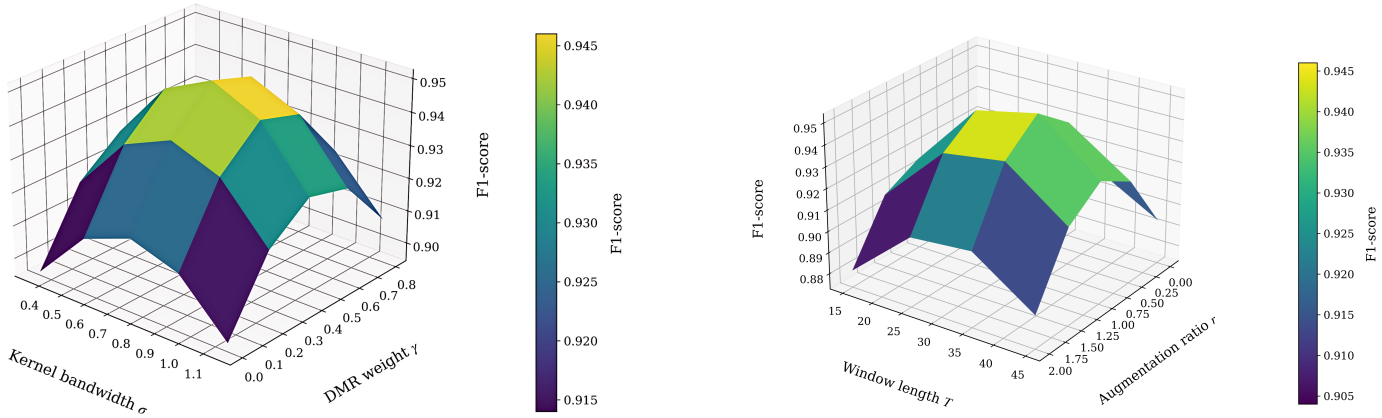


Figure 17. Joint hyperparameter interaction analysis. (**Left**): LKRP kernel bandwidth σ vs. DMR weight γ. (**Right**): CT–WGAN augmentation ratio r vs. window length T.

These results indicate that the proposed framework does not rely on fragile hyperparameter tuning, which is important for large-scale and long-term aviation monitoring deployments.

5.9.3. Feature Importance and Redundancy Suppression

To interpret the effect of LKRP with DMR regularization, Figure 3 visualizes the importance distribution and redundancy structure of the learned LKRP representation. The results show that the model emphasizes discriminative components while suppressing highly redundant ones, which is consistent with the ablation trend in Table 7, where removing DMR reduces both F1 and AUC [15,16]. This result highlights the importance of redundancy suppression for stable performance under nonstationary VHF conditions.

5.10. Real-Time Deployment Feasibility

In practical VHF monitoring, the detector must operate with low inference latency, modest computational cost, and stable performance during long-duration streaming.

Figure 18 compares inference latency across representative models, and Table 8 summarizes model size, FLOPs, and runtime. All evaluated methods satisfy the 100 ms real-time requirement, while the proposed framework achieves a favorable accuracy–latency balance: it improves detection performance (Table 3) with 0.69 M parameters, 0.71 M FLOPs, and 34.7 ms latency. Importantly, CT–WGAN and redundancy regularization are used only during training to mitigate class imbalance and improve feature quality; they are removed

during inference. The deployed system therefore consists only of the deterministic LKRP projection and a lightweight CNN–LSTM detector, yielding predictable runtime behavior compatible with real-time avionics constraints [7,12].

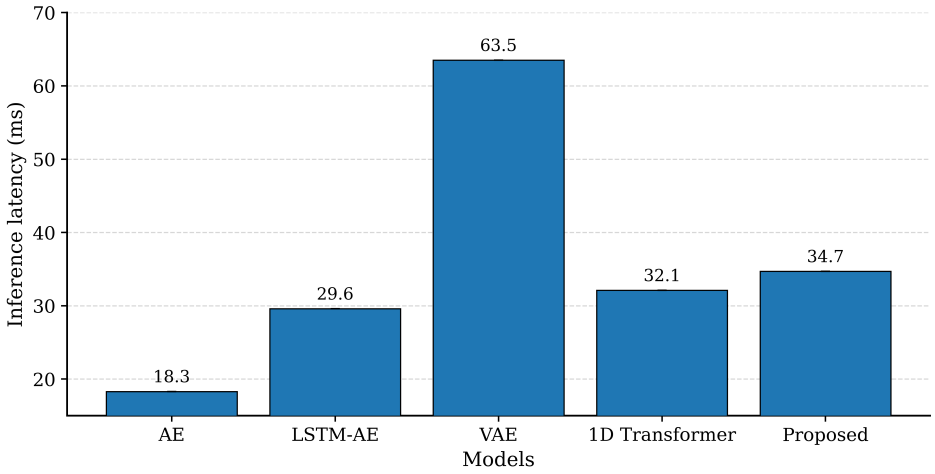


Figure 18. Inference latency comparison across models.

Table 8. Computational complexity and real-time latency comparison across models. Bold values indicate the proposed method.

Model	Params (M)	FLOPs (M)	Latency (ms)	Meets 100 ms?
Autoencoder (AE)	0.42	0.51	18.3	Yes
LSTM–AE	0.78	1.21	29.6	Yes
VAE	1.12	1.79	41.2	Yes
1D Transformer	1.85	3.46	63.5	Yes
CNN–LSTM (raw)	0.65	0.74	32.1	Yes
Proposed	**0.69**	**0.71**	**34.7**	**Yes**

5.11. Summary

Overall, the proposed LKRP + CT–WGAN + CNN–LSTM framework achieves the best detection performance among the evaluated methods while remaining compatible with real-time deployment constraints [7,12]. The ablation results indicate that CT–WGAN contributes substantially under extreme imbalance, and that LKRP and DMR provide complementary gains by improving nonlinear separability and suppressing redundant components. The latency and complexity results further show that, after removing training-only modules, the deployed LKRP projection with a lightweight CNN–LSTM detector offers a practical accuracy–latency balance for safety-critical VHF monitoring and PHM-enabled operational use.

6. Conclusions and Future Work

This study investigated real-time anomaly detection for civil aviation VHF communication, where nonstationarity, extreme anomaly sparsity, and strict latency constraints make conventional monitoring insufficient. The proposed LKRP–CT–WGAN–CNN–LSTM pipeline combines adaptive nonlinear representation learning, redundancy-aware feature shaping, and condition-consistent temporal augmentation, while maintaining lightweight inference.

The proposed method is designed as a monitoring tool rather than a replacement for existing communication systems. It could be integrated into ground or onboard monitoring platforms to support maintenance decisions. In safety-critical environments, stable behavior and low false-alarm rates remain essential.

On a dataset containing more than 1.2×10^6 VHF frames, the proposed framework achieves an F1-score of 0.947 ± 0.003 and an AUC of 0.972 ± 0.002 (averaged over five runs), while the average inference latency remains well below the 100 ms requirement. The ablation results show that LKRP improves the separability of weak degradation patterns, while conditional temporal augmentation increases the coverage of minority anomalies under severe class imbalance. During deployment, only the deterministic LKRP projection and the CNN–LSTM detector are required, which simplifies runtime behavior for long-term monitoring.

In deployment, the training and runtime stages are separated. Generative augmentation and redundancy regularization are used only during training, while the deployed system consists of the LKRP projection and a lightweight CNN–LSTM detector. This design supports stable runtime behavior during continuous monitoring.

The framework still has practical limitations. In particular, it relies on labeled anomaly samples and cannot fully address scenarios where annotations are unavailable. Generative augmentation also cannot ensure coverage of degradation modes that have not yet been observed. In addition, joint optimization of learnable kernels, mutual-information estimation, and temporal GAN modules introduces considerable training complexity.

Future work will explore strategies for reducing the dependence on labeled data, improving cross-domain generalization across communication environments, and incorporating simulation-driven monitoring with digital twin technologies. Further work will also explore model compression and hardware-aware optimization for deployment on resource-constrained airborne platforms. These results suggest that intelligent communication monitoring can become an important component of next-generation aviation PHM systems.

Author Contributions: Conceptualization, J.Z. and Y.H.; methodology, J.Z. and G.S.; software, J.Z.; validation, J.Z. and Z.L.; formal analysis, J.Z.; investigation, J.Z.; resources, Z.L. and Q.C.; data curation, Z.L. and Q.C.; writing—original draft preparation, J.Z.; writing—review and editing, Y.H. and G.S.; visualization, J.Z.; supervision, Y.H.; project administration, Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The operational VHF dataset used in this study cannot be publicly released due to aviation data confidentiality restrictions and regulatory considerations.

Acknowledgments: The authors thank the supporting institutions and collaborators who contributed to this work.

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

Abbreviations

VHF	Very High Frequency
PHM	Prognostics and Health Management
LKRP	Learnable Kernel–Redundancy Projection
DMR	differentiable minimum redundancy maximum relevance

CT–WGAN	conditional temporal Wasserstein GAN
CNN	convolutional neural network
LSTM	Long Short-Term Memory
SNR	signal-to-noise ratio
RSSI	received signal strength indicator
BER	bit error rate
PER	packet error rate
ROC	Receiver Operating Characteristic
PR	Precision–Recall
AUC	Area Under the Curve
FAR	false alarm rate

References

1. Haas, E. Aeronautical channel modeling. *IEEE Trans. Veh. Technol.* **2002**, *51*, 254–264. [CrossRef]
2. Zhou, M.; Kong, M.; Ye, Y.; Deng, B.; Tang, Y. Identifying sources of interference in civil aviation radio communication. *EURASIP J. Adv. Signal Process.* **2024**, *2024*, 88. [CrossRef]
3. Malik, A.; Rao, M. Radio frequency interference, its mitigation and its implications for the civil aviation industry. *Electronics* **2025**, *14*, 2483. [CrossRef]
4. Hosokawa, K.; Sakai, J.; Tomizawa, I.; Saito, S.; Tsugawa, T.; Nishioka, M.; Ishii, M. A monitoring network for anomalous propagation of aeronautical VHF radio waves due to sporadic E in Japan. *Earth Planets Space* **2020**, *72*, 88. [CrossRef]
5. Ain, Q.U.; Jilani, A.A.A.; Butt, N.A.; Rehman, S.U.; Alhulayyil, H.A. Anomaly detection for aviation cyber-physical system: Opportunities and challenges. *IEEE Access* **2024**, *12*, 175905–175925. [CrossRef]
6. Fix, J.; Ren, C.; Costa Lopes, A.; Morice, G.; Kobayashi, S.; Leterte, T.; Hinostroza Sáenz, I.D. Deep learning for aircraft classification from VHF radar signatures. *IET Radar Sonar Navig.* **2021**, *15*, 697–707. [CrossRef]
7. Edozie, E.; Shuaibu, A.N.; Sadiq, B.O.; John, U.K. Artificial intelligence advances in anomaly detection for telecom networks. *Artif. Intell. Rev.* **2025**, *58*, 100. [CrossRef]
8. Kohli, M.; Chhabra, I. A comprehensive survey on techniques, challenges, evaluation metrics and applications of deep learning models for anomaly detection. *Discov. Appl. Sci.* **2025**, *7*, 784. [CrossRef]
9. Schummer, P.; del Rio, A.; Serrano, J.; Jimenez, D.; Sánchez, G.; Llorente, Á. Machine learning-based network anomaly detection: Design, implementation, and evaluation. *AI* **2024**, *5*, 2967–2983. [CrossRef]
10. Fu, S.; Avdelidis, N.P. Prognostic and health management of critical aircraft systems and components: An overview. *Sensors* **2023**, *23*, 8124. [CrossRef]
11. Kordestani, M.; Orchard, M.E.; Khorasani, K.; Saif, M. An overview of the state of the art in aircraft prognostic and health management strategies. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3505215. [CrossRef]
12. Duraj, A.; Szczepaniak, P.S.; Sadok, A. Detection of anomalies in data streams using the LSTM-CNN model. *Sensors* **2025**, *25*, 1610. [CrossRef] [PubMed]
13. Ren, Z.; Jiang, Y.; Yang, X.; Tang, Y.; Zhang, W. Learnable faster kernel-PCA for nonlinear fault detection: Deep autoencoder-based realization. *J. Ind. Inf. Integr.* **2024**, *40*, 100622. [CrossRef]
14. Fan, J.; Chow, T.W.S.; Qin, S.J. Kernel-based statistical process monitoring and fault detection in the presence of missing data. *IEEE Trans. Ind. Inform.* **2021**, *18*, 4477–4487. [CrossRef]
15. Yu, K.; Li, W.; Xie, W.; Wang, L. A Hybrid Feature-Selection Method Based on mRMR and Binary Differential Evolution for Gene Selection. *Processes* **2024**, *12*, 313. [CrossRef]
16. Xie, S.; Zhang, Y.; Lv, D.; Chen, X.; Lu, J.; Liu, J. A new improved maximal relevance and minimal redundancy method based on feature subset. *J. Supercomput.* **2022**, *79*, 3157–3180. [CrossRef]
17. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680. [CrossRef]
18. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein GAN. *arXiv* **2017**, arXiv:1701.07875. [PubMed]
19. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A. Improved training of Wasserstein GANs. *arXiv* **2017**, arXiv:1704.00028. [CrossRef]
20. Yin, X.; He, W.; He, Q.; Chen, D.; Zhang, B.; Zhao, H. Data imbalanced fault diagnosis of aviation fuel pumps based on adaptive weighting using an enhanced broad learning system. *Measurement* **2026**, *257*, 118612. [CrossRef]
21. Zhou, Y.; Fu, C.; Wei, L.; Zhou, W.; Li, X.; You, Y. An integrated approach for addressing data imbalance in predicting fatality of helicopter accident. *Reliab. Eng. Syst. Saf.* **2025**, *259*, 111921. [CrossRef]

22. Baset, A.; Becker, C.; Derr, K.; Sarkar, S.; Kasera, S.K. Avisense: A real-time system for detection, classification, and analysis of aviation signals. *ACM Trans. Sens. Netw.* **2022**, *19*, 1–35. [CrossRef]
23. Yin, Z.; Zhang, F.; Yin, C.; Xu, G.; Liu, S. A bearing fault diagnosis method for sample imbalance. *Eng. Appl. Artif. Intell.* **2025**, *157*, 111171. [CrossRef]
24. Ma, M.; Han, L.; Zhou, C. BTAD: A benchmark for time-series anomaly detection and fault diagnosis in industrial applications. *Adv. Eng. Inform.* **2023**, *56*, 101949. [CrossRef]
25. Peng, C.; Hu, W.; Wang, L. Spectrum Anomaly Detection Based on Spatio-Temporal Network Prediction. *Electronics* **2022**, *11*, 1770. [CrossRef]
26. Dadhania, A.; Dave, P.; Bhatia, J.; Mehta, R.; Kumhar, M.; Tanwar, S.; Alabdulatif, A. Software defined network and graph neural network-based anomaly detection scheme for high speed networks. *Cyber Secur. Appl.* **2025**, *3*, 100079. [CrossRef]
27. Wang, J.; Huang, N.; Zhang, H.; Liu, L.; Fu, Q.; Cao, K.; Guo, X.; Jung, H. Self-learning model fusion for network anomaly detection: A hybrid CNN-LSTM-transformer framework. *PLoS ONE* **2025**, *20*, e0332502. [CrossRef]
28. Kaib, M.T.H.; Kouadri, A.; Harkat, M.F.; Bensmail, A.; Mansouri, M. Improving kernel PCA-based algorithm for fault detection in nonlinear industrial process through fractal dimension. *Process Saf. Environ. Prot.* **2023**, *179*, 525–536. [CrossRef]
29. Lin, H.; Huang, X.; Chen, Z.; He, G.; Xi, C.; Li, W. Matching Pursuit Network: An Interpretable Sparse Time-Frequency Representation Method Toward Mechanical Fault Diagnosis. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *36*, 12377–12388. [CrossRef]
30. Qi, J.; Chen, Z.; Kong, Y.; Qin, W.; Qin, Y. Attention-guided graph isomorphism learning: A multi-task framework for fault diagnosis and remaining useful life prediction. *Reliab. Eng. Syst. Saf.* **2025**, *263*, 111209. [CrossRef]
31. Gu, X.; Yu, Y.; Guo, L.; Gao, H.; Luo, M. CSWGAN-GP: A new method for bearing fault diagnosis under imbalanced condition. *Measurement* **2023**, *217*, 113014. [CrossRef]
32. Men, Z.; Li, Y.; Gao, L.; Zhang, Z. Fault diagnosis method for railway wagon bearings under imbalanced dataset based on improved ACWGAN. *Nonlinear Dyn.* **2025**, *113*, 14935–14962. [CrossRef]
33. Chen, L.; Li, Y.; Deng, X.; Liu, Z.; Lv, M.; Zhang, H. Dual Auto-Encoder GAN-Based Anomaly Detection for Industrial Control System. *Appl. Sci.* **2022**, *12*, 4986. [CrossRef]
34. Geiger, A.; Liu, D.; Alnegheimish, S.; Cuesta-Infante, A.; Veeramachaneni, K. TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks. *arXiv* **2020**, arXiv:2009.07769. [CrossRef]
35. Xu, J.; Wang, S.; Loscri, V.; Brighente, A.; Conti, M.; Rouvoy, R. GANSec: Enhancing Supervised Wireless Anomaly Detection Robustness through Tailored Conditional GAN Augmentation. *HAL* 2025, *preprint*. Available online: https://hal.science/hal-05137717v1 (accessed on 8 March 2026).
36. Qin, Z.; Zhang, Z.; Wu, Z.; Zhang, Q. Fault diagnosis method for rolling bearings under a small sample dataset based on improved WGAN-GP. *Measurement* **2025**, *237*, 118456. [CrossRef]
37. Belghazi, M.I.; Baratin, A.; Rajeshwar, S.; Ozair, S.; Bengio, Y.; Courville, A.; Hjelm, R.D. MINE: Mutual Information Neural Estimation. *arXiv* **2018**, arXiv:1801.04062.
38. Chen, L.; Liu, J. Identification of Shipborne VHF Radio Based on Deep Learning with Feature Extraction. *J. Mar. Sci. Eng.* **2024**, *12*, 810. [CrossRef]
39. Zahid, M.U.; Akram, M.U.; Nisar, M.D.; Maqsood, F.; Ali, S.U.; Montaha, M. CommRad RF: A dataset of communication radio signals for detection, identification and classification. *Data Brief* **2025**, *59*, 111387. [CrossRef]
40. Kumar, A.S.; Raja, S.; Pritha, N.; Raviraj, H.; Lincy, R.B.; Rubia, J.J. An adaptive transformer model for anomaly detection in wireless sensor networks in real-time. *Meas. Sens.* **2023**, *25*, 100625. [CrossRef]
41. Latif-Martínez, H.; Suárez-Varela, J.; Cabellos-Aparicio, A.; Barlet-Ros, P. GAT-AD: Graph Attention Networks for contextual anomaly detection in network monitoring. *Comput. Ind. Eng.* **2025**, *200*, 110830. [CrossRef]

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