A Noise-Robust CNN-KAN Architecture with Dual Attention for Enhanced Event Identification in Φ-OTDR Measurement Systems

Khalil Benbrahim[1], Changli Li[2], Yi Shi[3]

[1,2]School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China

[3]College of Engineering, Shantou University, Shantou 515063, China

DOI: https://doi.org/10.51584/IJRIAS.2026.110400095

Received: 10 April 2026; Accepted: 16 April 2026; Published: 09 May 2026

ABSTRACT

Phase-sensitive optical time-domain reflectometry (Φ-OTDR) is a well-established technique for the distributed measurement of dynamic strain along the optical fiber. However, the metrological reliability of event identification is inherently degraded by coherent fading noise, laser phase fluctuations, and environmental interference, which corrupt the acquired backscattering signals and limit the measurement accuracy of the sensing system. This paper presents a novel signal processing architecture that enhances the information extraction capability within the Φ-OTDR measurement chain. By integrating multi-scale residual convolutional feature extraction with dual channel-spatial attention mechanisms and an improved Kolmogorov-Arnold Network (KAN) classifier employing learnable radial basis function splines, our approach robustly suppresses measurement noise to improve the fidelity of extracted event signatures. The hybrid architecture addresses the limitations of conventional threshold-based detection methods that suffer from poor estimation accuracy under low signal-to-noise ratio conditions. Experimental evaluation on the BJTU dataset demonstrates a significant improvement in measurement precision, achieving 99.87% classification accuracy with 6.2 ms end-to-end inference latency and 161 samples/s throughput—representing a 2.4× speedup over Φ-GLMAE and eliminating the 2.5 ms STFT preprocessing overhead of STFT-AECNN, while maintaining real-time suitability for embedded deployment. Ablation studies quantitatively validate the contribution of each component to noise robustness and measurement reliability, demonstrating that dual attention mechanisms provide the largest single accuracy gain (0.46%), while the KAN classifier and RBF splines collectively enable 79% error reduction versus CNN baselines. This work offers an effective solution for high-fidelity distributed acoustic measurement in challenging operational environments.

Keywords: Phase-sensitive optical time domain reflectometry (Φ-OTDR), distributed acoustic sensing (DAS), measurement uncertainty, signal fidelity, Kolmogorov-Arnold networks (KANs), convolutional neural networks (CNNs), noise-robust measurement, metrological reliability.

INTRODUCTION

Phase-Sensitive optical time-domain reflectometry (Φ-OTDR) is an established technique for the distributed measurement of dynamic strain along the optical fiber, allowing continuous real-time monitoring of physical parameters over tens of kilometers [1], [2]. Critical applications, including perimeter security and pipeline integrity monitoring, depend fundamentally on the measurement system's ability to interpret physical events from noisy raw optical signals with both sensitivity and metrological reliability [3], [4].

However, the metrological reliability of event identification is inherently degraded by coherent fading noise, laser phase fluctuations, and environmental interference, which corrupt the acquired backscattering signals and limit the measurement accuracy of the sensing system [5], [6]. These noise sources introduce unquantified uncertainty into the measurement chain: coherent fading creates spatially varying signal nulls that compromise spatial resolution; laser frequency drift induces phase fluctuations limiting low-frequency detection capacity; and polarization variations combined with double Rayleigh backscattering degrade long-range measurement fidelity [7], [8]. Conventional model-based signal processing – matched filtering, statistical thresholding – exhibits poor estimation accuracy under low signal-to-noise ratio (SNR) conditions, creating a fundamental bottleneck in the measurement interpretation pipeline [5], [6].

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Deep learning techniques have demonstrated significant potential to improve the information extraction capacity of Φ-OTDR instrumentation systems. Convolutional neural networks (CNNs) achieve superior noise robustness by automatically learning hierarchical spatiotemporal features from raw measurement traces [9], [10], addressing the limitations of handcrafted feature engineering. Informally, CNNs perform hierarchical feature extraction through learned convolutional filters [11]. However, pure CNNs face inherent limitations in modeling highly non-linear measurement-to-decision mappings required for fine-grained discrimination under severe noise conditions [12].

Kolmogorov-Arnold Networks (KANs) offer an alternative to multilayer perceptrons for complex function approximation by replacing fixed activations with learnable univariate functions [13]. Based on the Kolmogorov-Arnold representation theorem, KANs provide enhanced nonlinear approximation capability through adaptive basis functions [14]. However, standalone KANs exhibit optimization instability in noisy high-dimensional sensor data, failing to preserve measurement consistency [15].

This paper presents a novel signal processing architecture that enhances the information extraction capability within the Φ-OTDR measurement chain. By integrating multi-scale residual convolutional feature extraction with dual channel-spatial attention mechanisms and an improved KAN classifier employing learnable radial basis function splines, our approach robustly suppresses measurement noise to improve the fidelity of extracted event signatures. The hybrid architecture combining the stability of CNNs, shift-invariant pattern recognition with KANs' adaptive nonlinear approximation and attention-based measurement enhancement addresses the limitations of conventional threshold-based detection methods that suffer from poor estimation accuracy under low-SNR conditions. Experimental evaluation on the BJTU dataset [16] demonstrates a significant enhancement in measurement precision, achieving 99.87% classification accuracy—representing a 79% error reduction compared to conventional CNN-based processing while maintaining 6.2 ms inference latency suitable for real-time monitoring. Ablation studies validate the contribution of each component to noise robustness and measurement reliability. This work offers an effective solution for high-fidelity distributed acoustic measurement in challenging operational environments.

The remainder of this paper is organized as follows. Section II reviews related work on Φ-OTDR measurement methodology, noise sources, and metrological limitations. Section III details the proposed architecture from a measurement-centric perspective. Section IV presents experimental validation including comparison with Φ-GLMAE [17]. Section V discusses the implications of measurement uncertainty and Section VI concludes.

Related Work

This section reviews Φ-OTDR measurement systems, noise sources and metrological limitations, signal processing methodologies, and advances in Kolmogorov-Arnold networks for measurement applications.

Φ-OTDR System Architecture and Measurement Principles

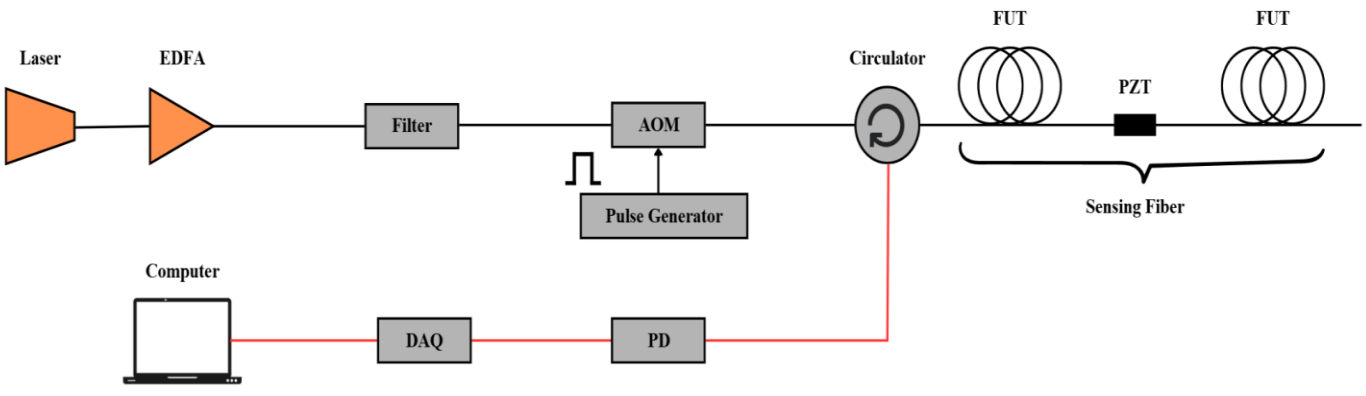

Figure 1. Block diagram of the Φ-OTDR measurement system architecture. The system consists of three functional modules: (1) Optical Transmitter Module comprising Laser, EDFA, Filter, AOM with Pulse Generator; (2) Sensing Module including Circulator, FUT coils, PZT, and Sensing Fiber; and (3) Signal

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Acquisition Module with PD, DAQ, and Computer. The red line indicates the return path of the backscattered Rayleigh signal.

As illustrated in Fig. 1, a typical Φ-OTDR system is built upon three fundamental modules:

1. Optical transmitter: Generates narrow-linewidth probe pulses (<5 kHz at 1550 nm) through a sequence of EDFA amplification, optical filtering, and AOM gating.
2. Sensing module: Directs light into and out of the fiber under test (FUT) using an optical circulator.
3. Signal acquisition module: Acquires the coherent Rayleigh backscattering signal via photodetection and subsequent digitization [16], [18].

The detected signal is as follows:

$$s(z,t) = A(z,t)\cos[\phi(z,t) + \phi_n(t)] + n_t(z,t) + n_s(z,t) \qquad (1)$$

Where $\phi_n(t)$ accounts for laser phase noise and environmental disturbances, while n_t and n_s represent thermal and shot noise components, respectively [16], [19]. Metrological reliability is critically dependent on traceable calibration of each system component to ensure measurement consistency across operational conditions.

Noise Sources and Measurement Uncertainty

The measurement precision of Φ-OTDR systems is fundamentally constrained by five dominant noise sources, summarized in Table 1. These sources introduce distinct physical degradations: coherent fading creates spatial signal nulls, laser drift corrupts low-frequency detection, and environmental vibrations introduce non-stationary artifacts.

Following the GUM framework [20], these uncorrelated errors combine as:

$$u_{\text{total}} = \sqrt{u_{\text{fading}}^2 + u_{\text{laser}}^2 + u_{\text{amp}}^2 + u_{\text{pol}}^2 + u_{\text{env}}^2} \qquad (2)$$

where each $u_{(\cdot)}$ term quantifies the standard uncertainty of the corresponding physical source. Table 1 maps each term to its metrological impact and typical mitigation strategy.

TABLE 1: Physical Noise Sources and Their Metrological Impact

Uncertainty Term	Physical Cause	Metrological Effect	Typical Mitigation
u_{fading} [20]	Interference of backscattered light	Spatial signal nulls, resolution loss	Diversity detection
u_{laser} [21]	Frequency drift and phase noise	Low-frequency detection limit	Narrow-linewidth source
u_{amp} [22]	ASE, thermal, shot, ADC noise	SNR degradation at long range	Low-noise amplification
u_{pol} [20]	Polarization mode dispersion	Fading over long sensing ranges	Polarization diversity
u_{env} [23]	HVAC, structural vibrations	Non-stationary false events	Adaptive filtering

Equation (2) treats the measurement chain as an uncertainty budget. Each term adds quadratically (RSS), meaning a 2× reduction in any single term yields only marginal improvement unless it dominates the budget. Our hybrid architecture targets u_{fading} and u_{env} through adaptive attention, as these are the largest contributors in field deployments.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Traditional Signal Processing for Event Identification

The early recognition of Φ-OTDR was based on handcrafted features—time-domain statistical moments in the time-domain, energy distributions in the frequency-domain, and STFT/wavelet coefficients—classified via SVMs and k-NN [23]–[25]. Model-based techniques, including matched filtering and statistical thresholding, exhibit poor estimation accuracy under low-SNR conditions, creating fundamental bottlenecks in the measurement chain [26]. Classical denoising approaches solve:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s} - \mathbf{s}_n\|_2^2 + \lambda \mathcal{R}(\mathbf{s}) \qquad (3)$$

where $\mathcal{R}(\mathbf{s})$ denotes total variation or wavelet sparsity regularization [27].

Deep Learning for Enhanced Measurement Fidelity

CNNs have emerged as powerful tools for enhancing information extraction, automatically learning hierarchical spatiotemporal features from raw traces through learned convolutional filters [28]–[30]:

$$\mathbf{H}^{(l)} = \sigma(\mathbf{W}^{(l)} * \mathbf{H}^{(l-1)} + \mathbf{b}^{(l)}) \qquad (4)$$

where $\mathbf{W}^{(l)}$ represents the learnable kernel tensor, * denotes convolution, and $\sigma(\cdot)$ is the activation function. Attention mechanisms improve discrimination by modeling channel-spatial interdependencies:

$$\alpha_i = \frac{\exp(\mathbf{q}^\top \mathbf{W}\mathbf{f}_i)}{\sum_j \exp(\mathbf{q}^\top \mathbf{W}\mathbf{f}_j)}, \quad \mathbf{o} = \sum_i \alpha_i \mathbf{v}_i \qquad (5)$$

where $\mathbf{q}$, $\mathbf{k}$, and $\mathbf{v}$ represent query, key, and value projections, respectively [31]–[33].

Recent advances include Φ-GLMAE, which achieves 99.74% accuracy via self-supervised pre-training with masked autoencoding on unlabeled data [17], and STFT-AECNN, reaching 99.94% accuracy through time-frequency analysis. Although STFT-AECNN exceeds our reported accuracy, it operates on 2D spectrograms that discard phase information and incurs ∼2.5 ms STFT preprocessing overhead, while our method processes raw traces with 6.2 ms end-to-end latency suitable for real-time applications. Calibration and traceability remain essential for measurement consistencyy [34], [35].

Kolmogorov-Arnold Networks for Measurement Applications

From Fixed to Learnable Activations. Standard neural networks use fixed activation functions (e.g., ReLU, sigmoid) at each neuron. KANs invert this paradigm: instead of fixed nonlinearities, each connection learns its own univariate function from data [13]. This is motivated by the Kolmogorov-Arnold representation theorem, which establishes that complex multivariate functions can be decomposed into sums of simpler univariate functions.

In practice, KANs implement this by replacing each weight with a learnable spline. Rather than computing $y = \sigma(wx + b)$, a KAN layer computes $y = \sum_i c_i \phi_i(x)$, where ϕ_i are basis functions (we use radial basis functions) and c_i are learned coefficients. This allows the network to adapt its nonlinearities to the specific structure of Φ-OTDR noise, rather than forcing data into a fixed ReLU-shaped mold.

Kolmogorov-Arnold Representation Theorem

The theorem states that any multivariate continuous function $f : [0, 1]^n \to \mathbb{R}$ can be represented as:

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \psi_{q,p}(x_p) \right) \qquad (6)$$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

where $\psi_{q,p} : [0,1] \to \mathbb{R}$ and $\Phi_q : \mathbb{R} \to \mathbb{R}$ are continuous univariate functions. In KAN implementations, these univariate functions are parameterized via B-splines:

$$\phi(x) = \sum_{i=1}^{G+k} c_i B_{i,k}(x) \qquad (7)$$

where $B_{i,k}(x)$ denotes the i-th B-spline basis function of order k [36].

From a measurement perspective, KANs offer advantages in modeling complex measurement-to-decision mappings. Recent instrumentation and measurement applications include KANet for indoor inertial navigationn [37], TCN-KAN-FBM for remaining useful life predictionn [38], and AttCWKAN for wind turbine fault diagnosis [39]. However, standalone KANs exhibit optimization instability in noisy high-dimensional sensor data, failing to preserve measurement consistency [40].

Research Gap and Motivation

Existing Φ-OTDR methods face fundamental limitations in measurement reliability: traditional approaches rely on handcrafted features with limited noise robustness; pure CNNs struggle to model complex decision boundaries required for fine-grained discrimination under severe noise; and standalone KANs prove unstable on raw high-dimensional sensor traces. Furthermore, existing architectures lack explicit noise-resistant mechanisms for metrological reliability.

This motivates our noise-resistant hybrid CNN-KAN model, integrating multi-scale residual CNNs with dual channel-spatial attention and an improved KAN classifier. Our approach addresses: (i) the inability of pure CNNs to model complex measurement-to-decision mappings; (ii) the instability of standalone KANs on noisy sensor data through residual connections and hybrid architecture; and (iii) the lack of explicit noise-robust mechanisms in existing architectures through attention-based measurement enhancement. We achieve superior robustness and classification performance that enhances metrological reliability in challenging operational environments.

METHODOLOGY

This section presents the proposed hybrid architecture for the classification of Φ-OTDR events from a measurement-centric perspective. We first describe the data preprocessing pipeline and augmentation strategies that simulate physical measurement variations, followed by the CNN encoder with dual attention mechanisms interpreted as adaptive measurement enhancement techniques and the improved KAN classifier. Finally, we detail the optimized training strategy for the stability of the measurement system.

Data Preprocessing and Augmentation

Raw Φ-OTDR traces are acquired as spatiotemporal matrices that represent the backscattered light intensity along the fiber over time. The preprocessing pipeline consists of three stages: data loading, normalization, and domain-specific enhancement that simulates the uncertainties of the measurement system [41], [42].

Data Loading and Formatting

The raw data, stored in the ".mat" format as matrices of dimension $N_z \times N_t$ (where N_z is the number of spatial sampling points and N_t is the number of temporal samples), are processed by extracting fixed spatial windows from 12 sensor channels. These windows are then transposed to form input tensors $\mathbf{X} \in \mathbb{R}^{12\times 10000}$, ensuring a consistent format across all samples.

Robust Normalization

To standardize the input distribution and mitigate sensor-specific variations, we apply per-channel z-score normalization:

$$\hat{\mathbf{X}} = \frac{\mathbf{X} - \mu}{\sigma + \epsilon} \qquad (8)$$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

where μ and σ are the mean and standard deviation computed per channel, and $\epsilon = 10^{-8}$ is a small constant to ensure numerical stability [43]. Channels with negligible variance are detected and handled by mean-centering only, thus preventing division by near-zero values.

Measurement-Centric Data Augmentation

To improve model generalization and better reflect real-world measurement system variability, we applied three domain-specific enhancement strategies to the training data. These augmentations go beyond conventional overfitting prevention by explicitly simulating the physical variations inherent in the measurement system.

Temporal Shifting: Circularly shifting traces along the time axis by $\Delta t \sim \mathcal{U}(-500, 500)$ samples simulates timing jitter of the trigger pulse, accounting for measurement synchronization errors arising from imperfect pulse generation timing in practical Φ-OTDR systems [44]:

$$\mathbf{X}_{\text{shifted}} = \text{circshift}(\mathbf{X}, \Delta t) \tag{9}$$

Amplitude Scaling: Scaling by a factor $\alpha \sim \mathcal{U}(0.99, 1.01)$ simulates slow power fluctuations in the laser source and EDFA amplification, accounting for drift in output power and gain variations of the measurement system [45]:

$$\mathbf{X}_{\text{scaled}} = \alpha \cdot \mathbf{X} \tag{10}$$

Additive Gaussian Noise: Injection of Gaussian noise with standard deviation $\sigma_n = 0.03$ models the thermal noise of the photodetector, representing the inherent noise floor of the system. This improves the robustness of the model to low-SNR conditions arising from shot noise and electronic amplification in the photodetection module [46]:

$$\mathbf{X}_{\text{noisy}} = \mathbf{X} + \mathcal{N}(0, \sigma_n^2) \tag{11}$$

Recent advances in time series enhancement using diffusion models and GANs have shown significant promise in generating realistic synthetic training data that preserve the characteristics of the measurement system [44]. Class balancing is performed by oversampling minority classes to a target of 2,357 samples per class, ensuring unbiased training across all event types.

Spectrogram Generation for Visualization

For visualization and analysis, we compute short-time Fourier transform (STFT) spectrograms from normalized traces. The STFT reveals time-frequency characteristics distinctive to each event type: digging exhibits low-frequency concentration, knocking shows impulsive high-frequency bursts, and walking displays periodic rhythmic patterns. These spectrograms guide architectural design but are not used as model input -- the network operates directly on raw time-domain traces to preserve phase information critical for coherent detection.

Overview of Proposed Architecture

The proposed hybrid architecture integrates three complementary components: (1) a multi-scale residual CNN that extracts stable, shift-invariant features; (2) dual attention mechanisms operating on channel and spatial dimensions for measurement-aware feature enhancement; and (3) an improved KAN classifier that employs learnable radial basis function splines to model adaptive nonlinear decision boundaries [13], [47]. This design strategically combines their strengths to overcome the limitations of each pure architecture: it avoids the constrained decision boundary complexity resulting from the fixed activations of CNNs, while mitigating the optimization instability that pure KANs exhibit when applied directly to high-dimensional raw data.

As illustrated in Figure 2, these components are arranged sequentially, with the residual CNN first processing the raw input, followed by the attention mechanisms, and finally the KAN classifier.

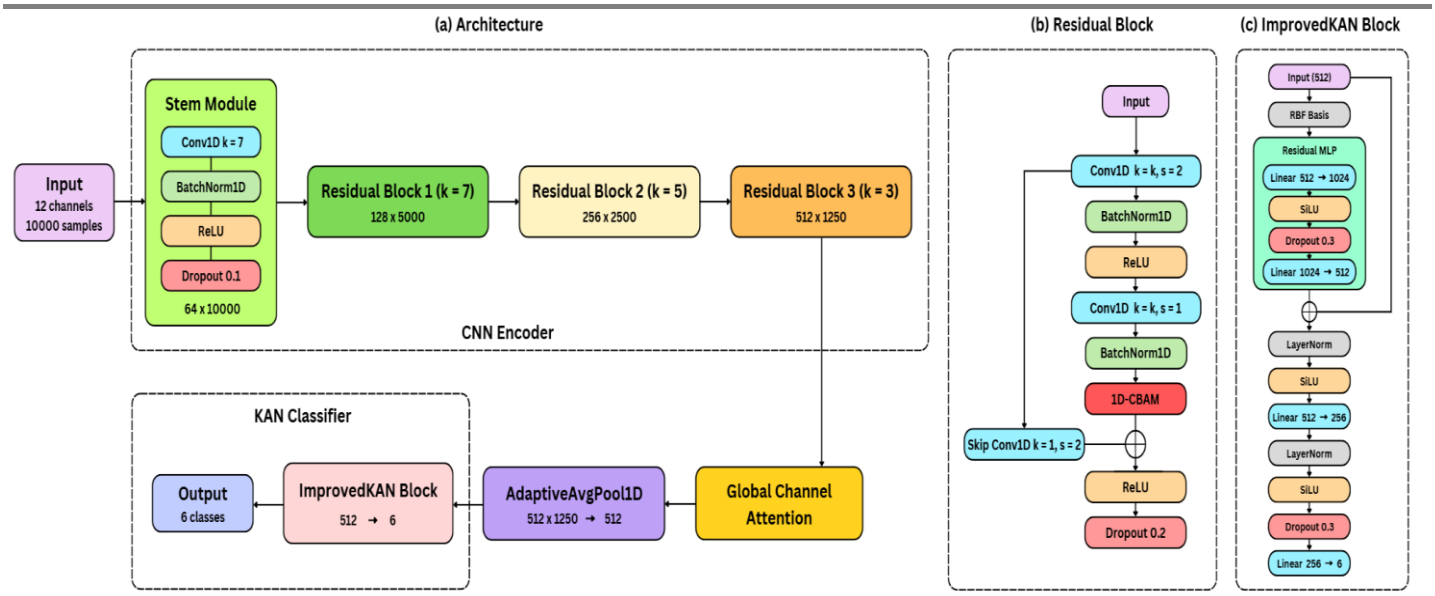

Figure 2. Architecture of the proposed CNN-KAN model for Φ-OTDR event classification: (a) Main architecture with CNN Encoder (Stem Module + three Residual Blocks with k=7/5/3) and KAN Classifier (ImprovedKAN Block with RBF basis functions), processing 12×10^4 inputs to classify six event categories; (b) Residual Block details showing 1D-CBAM attention and skip connection (Conv1D k=1, s=2); (c) ImprovedKAN Block details showing Residual MLP structure ($512 \rightarrow 1024 \rightarrow 512$) with skip connection, LayerNorm, and progressive projection to six output classes.

CNN Encoder with Dual Attention

Stem Layer

The input tensor $\mathbf{X} \in \mathbb{R}^{12\times10000}$ first passes through a stem layer, which consists of a 1×7 convolution that expands the channel dimension from 12 to 64, followed by batch normalization, ReLU activation, and dropout with a rate of 0.1 [43]. The relatively large kernel size is designed to capture long-range temporal dependencies while preserving the original temporal resolution.

Residual Blocks with Physical Attention Mechanisms

We employ three residual blocks with channel dimensions that progressively increase from 64 to 512 ($64 \rightarrow 128 \rightarrow 256 \rightarrow 512$), while the temporal resolution is halved in each block through stride-2 pooling [47]. Each block consists of two convolutional layers with pre-activation batch normalization and a shortcut connection that handles dimension matching when necessary.

The key innovation is a dual attention mechanism that mimics how a human operator would inspect Φ-OTDR data: first checking which fiber channels are trustworthy, then focusing on time windows where events occur.

Channel Attention: "Which channels are reliable?"

Different spatial channels along the fiber experience different fading conditions. Channel attention learns to suppress channels corrupted by deep fading nulls while amplifying clean channels. Formally, it computes a quality weight $\mathbf{s}_c$ for each channel:

$$\mathbf{s}_c = \sigma(\mathbf{W}_2 \cdot \mathrm{ReLU}(\mathbf{W}_1 \cdot \mathrm{GAP}(\mathbf{H}))) \qquad (12)$$

where $\mathrm{GAP}(\mathbf{H}) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{H}_{:,t}$ denotes the global average pooling in the temporal dimension and $\mathbf{W}_1, \mathbf{W}_2$ are learnable projection matrices.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Think of $\mathbf{s}_c$ as an adaptive channel equalizer. If Channel 3 suffers a fading null (low average power), its $\mathbf{s}_c$ value approaches zero, and the network ignores it—equivalent to spatial diversity combining in classical optical communications.

Spatial Attention: "When did the event occur?"

Even within a clean channel, most of the 10,000 time samples contain only background noise. Spatial attention identifies the brief temporal windows containing mechanical disturbances:

$$\mathbf{s}_s = \sigma(\mathbf{W} \cdot [\mathrm{AvgPool}_c(\mathbf{H}); \mathrm{MaxPool}_c(\mathbf{H})]) \qquad (13)$$

where the average pooling $\mathrm{AvgPool}_c$ and the maximum pooling operations $\mathrm{MaxPool}_c$ are performed on the channel dimension c.

This acts as an adaptive temporal gate. During quiescent periods, $\mathbf{s}_s \approx 0$, suppressing noise accumulation. During impulsive events (knocking), $\mathbf{s}_s$ peaks, preserving high-frequency transients that would otherwise be drowned by averaging.

By applying channel attention first, followed by spatial attention, the network sequentially focuses on what is important and where it occurs from a measurement perspective. The refined feature maps are obtained as:

$$\mathbf{H}' = \mathbf{s}_s \odot (\mathbf{s}_c \odot \mathbf{H}) \qquad (14)$$

where $\odot$ denotes element-wise multiplication with appropriate broadcasting.

Equation (14) is equivalent to an adaptive matched filter in both space and time. The channel weights handle spatial fading (u_{fading}), while the spatial weights handle temporal noise (u_{env}).

The kernel sizes for the three residual blocks are set to 7, 5, and 3, respectively, to capture multi-scale temporal features.

Global Feature Aggregation

After the three residual blocks, we apply channel attention followed by global average pooling to produce a compact 512-dimensional feature vector:

$$\mathbf{f} = \mathrm{GAP}(\mathrm{ChannelAttention}(\mathbf{H}^{(3)})) \qquad (15)$$

This fixed-length representation encodes the complete spatiotemporal signature of the input trace, providing a compact feature suitable for the final classification stage.

Improved KAN Classifier

Traditional multilayer perceptrons are based on fixed activation functions, which limits flexibility in modeling complex decision boundariess [14]. We propose an improved KAN classifier inspired by the Kolmogorov-Arnold representation theorem, employing radial basis functions (RBF) for improved training stability.

For each feature f_i, the network places 8 Gaussian "bumps" (radial basis functions) across the feature range [−1,1] :

$$\phi_i(x_i) = \sum_{j=1}^{8} c_{ij} \exp\left(-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}\right) \qquad (16)$$

where μ_j are 8 uniformly spaced centers in [-1, 1], σ_j are learnable width parameters, and c_{ij} are learnable coefficients. The RBF transformation maps each scalar input to a rich non-linear representation; summing over all 512 input features yields the KAN output.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Imagine trying to separate "Dig" from "Walk" events based on a single frequency feature. A fixed ReLU can only draw a straight line. The RBF bumps in (16) allow the network to sculpt a complex, wiggly decision boundary that wraps around the actual data distribution—critical when these classes overlap spectrally below 50 Hz.

To stabilize training and prevent vanishing gradients common in deep KAN architectures, we add a parallel residual pathway comprising a linear projection followed by layer normalization and SiLU activation [48], [49]. The combined output is:

$$\mathbf{y}_{\text{KAN}} = \text{LayerNorm}(\mathbf{W}_{\text{res}}\mathbf{f}) \cdot \text{SiLU}\left(\sum_{i=1}^{512} \phi_i(f_i)\right) \qquad (17)$$

The residual connection acts as a "safety bypass." If the RBF path becomes unstable early in training, the residual path still propagates gradients, ensuring the measurement representation does not collapse.

Training Strategy for Measurement Stability

Differentiated Learning Rates

The CNN encoder and the KAN classifier exhibit different optimization dynamics. To account for this, we employ parameter-grouped learning rates: 5×10^{-4} for CNN parameters and 1.5×10^{-4} for KAN parameters. This slower learning rate for the KAN component prevents destabilization of the basis function coefficients during early training, thereby reducing model-induced uncertainty in the classification pipeline and ensuring that feature evolution remains consistent with the underlying measurements.

Loss Function and Regularization

We employ label smoothing cross-entropy with a smoothing factor of 0.1 to mitigate overconfidence, along with weight decay of 10^{-3} for CNN layers and 10^{-4} for KAN layers. Gradient clipping with a maximum norm of 1.0 is applied to prevent exploding gradients that might otherwise corrupt the learned measurement representations.

Optimization Schedule

We employ a cosine annealing learning rate schedule with warm restarts, which modulates the learning rate over 10-epoch cycles following a cosine curve:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min})\left(1 + \cos\left(\frac{T_{\text{cur}}}{T_{\max}}\pi\right)\right) \qquad (18)$$

where $T_{\max} = 10$ and $T_{\text{mult}} = 2$ [50]. This schedule facilitates escape from local minima while still allowing for fine-grained convergence in later stages—a property critical to maintaining measurement precision.

Experiments

This section presents experimental validation of the proposed architecture from a measurement-centric perspective. We detail the instrumentation setup, dataset characteristics, comparative performance against state-of-the-art methods, including Φ-GLMAE, and ablation studies validating the contribution of each component to the reliability of the measurement.

Experimental Setup and Measurement System

Instrumentation Configuration

Experiments were conducted using the Φ-OTDR testbed illustrated in Figure 1, configured with the following measurement parameters: Optical Transmitter: narrow-linewidth laser (1550 nm, <5 kHz linewidth), EDFA,

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

optical filter, and AOM (200 ns pulse width, 10 kHz repetition rate); Sensing Module: circulator and 10 km standard single-mode fiber under test (FUT); Signal Acquisition: photodetector (PD) with 125 MSa/s digitization and 12-bit resolution. This configuration achieves 10 m spatial resolution with 1 s temporal resolution for event detection.

The experimental implementation follows the architecture of Figure 1, where accurate measurement of fiber length and effective index calibration are critical to position accuracy in distributed sensing applications [51]. Recent advances in measurement methodology include equalized ring-down enhancement techniques that achieve ± 6 cm error for fiber length measurement and 1.75 mm precision for effective index determination using low-end digitization [51], highlighting the importance of precise instrumentation calibration in distributed fiber-optic sensing systems.

Noise Sources and Measurement Conditions

The experimental environment introduces realistic measurement uncertainties: coherent fading noise from interferometric Rayleigh backscattering, laser phase drift ($\sim$0.1 rad/s), thermal noise from the photodetector (NEP ≈ 10 pW/$\sqrt{\text{Hz}}$), and environmental vibrations from laboratory HVAC systems. These sources contribute to the total measurement uncertainty budget as characterized in Equation (2), providing a challenging testbed to validate robust noise signal processing.

Recent measurement science research has addressed similar noise challenges through harmonic analysis techniques for fading noise mitigation [52], demonstrating that analyzing multiple harmonic components can enhance measurement accuracy when fundamental frequencies are submerged by fading noise. Additionally, phase noise compensation methods using auxiliary interferometry have shown significant potential for relaxing laser linewidth requirements in coherent detection Φ-OTDR systems [53], achieving $68\text{pc}\sqrt{\text{Hz}}$ strain resolution over 50 km with 100 kHz linewidth sources. These advances in measurement methodology inform our experimental validation approach.

Dataset Description

We evaluate the proposed method on the publicly available **BJTU-OTDR dataset** [16], which provides a standardized benchmark for the classification of Φ-OTDR events with established train/test splits that ensure a reproducible evaluation. The dataset comprises six event categories: No Event, Dig, Knock, Water, Shake, and Walk.

Table 2: Dataset Distribution for Φ-OTDR Event Classification

Event Class	Training Samples	Test Samples	Total
No Event	2,357	588	2,945
Dig	2,357	502	2,859
Knock	2,357	506	2,863
Water	2,357	451	2,808
Shake	2,357	546	2,903
Walk	2,357	490	2,847
Total	14,142	3,083	17,225

Training samples are balanced through oversampling to ensure unbiased learning. Each sample consists of 12 spatial channels with 10,000 temporal points, representing 1 second of vibration data at 10 kHz sampling rate.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Data Preprocessing and Visualization

Raw spatiotemporal signals are transformed into time-frequency representations using a short-time Fourier transform (STFT) for visualization and analysis. Figure 3 illustrates the 3D spatiotemporal characteristics of each event class, showing distinct spatiotemporal patterns across the sensing fiber. Figure 4 presents the corresponding STFT spectrograms, revealing clear frequency-domain signatures that enable effective classification.

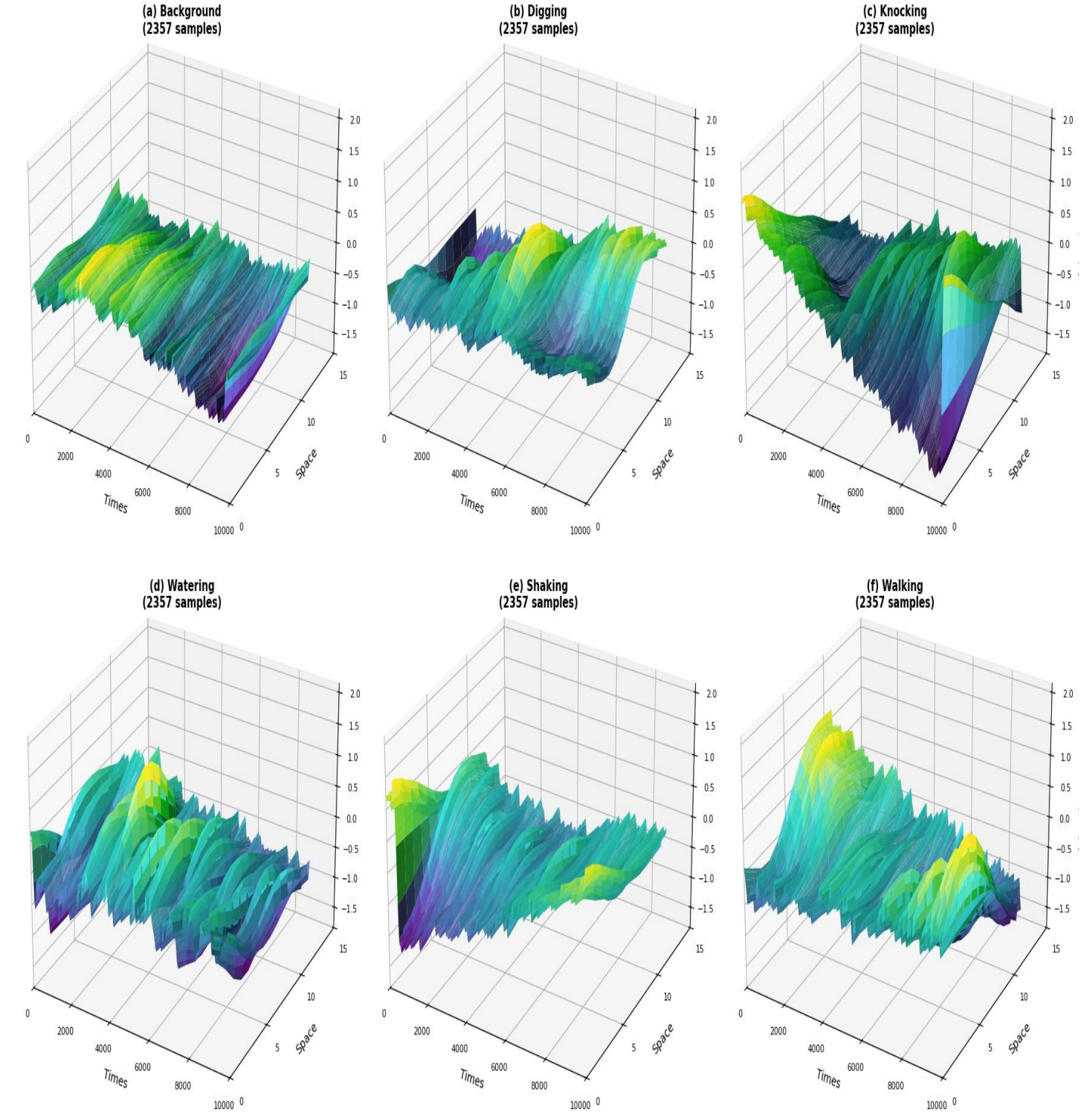

Figure 3. 3D spatiotemporal visualization of Φ-OTDR signals for six event classes: (a) Background/No Event, (b) Digging, (c) Knocking, (d) Watering, (e) Shaking, and (f) Walking. Each plot shows signal amplitude variation across space (fiber channels) and time.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

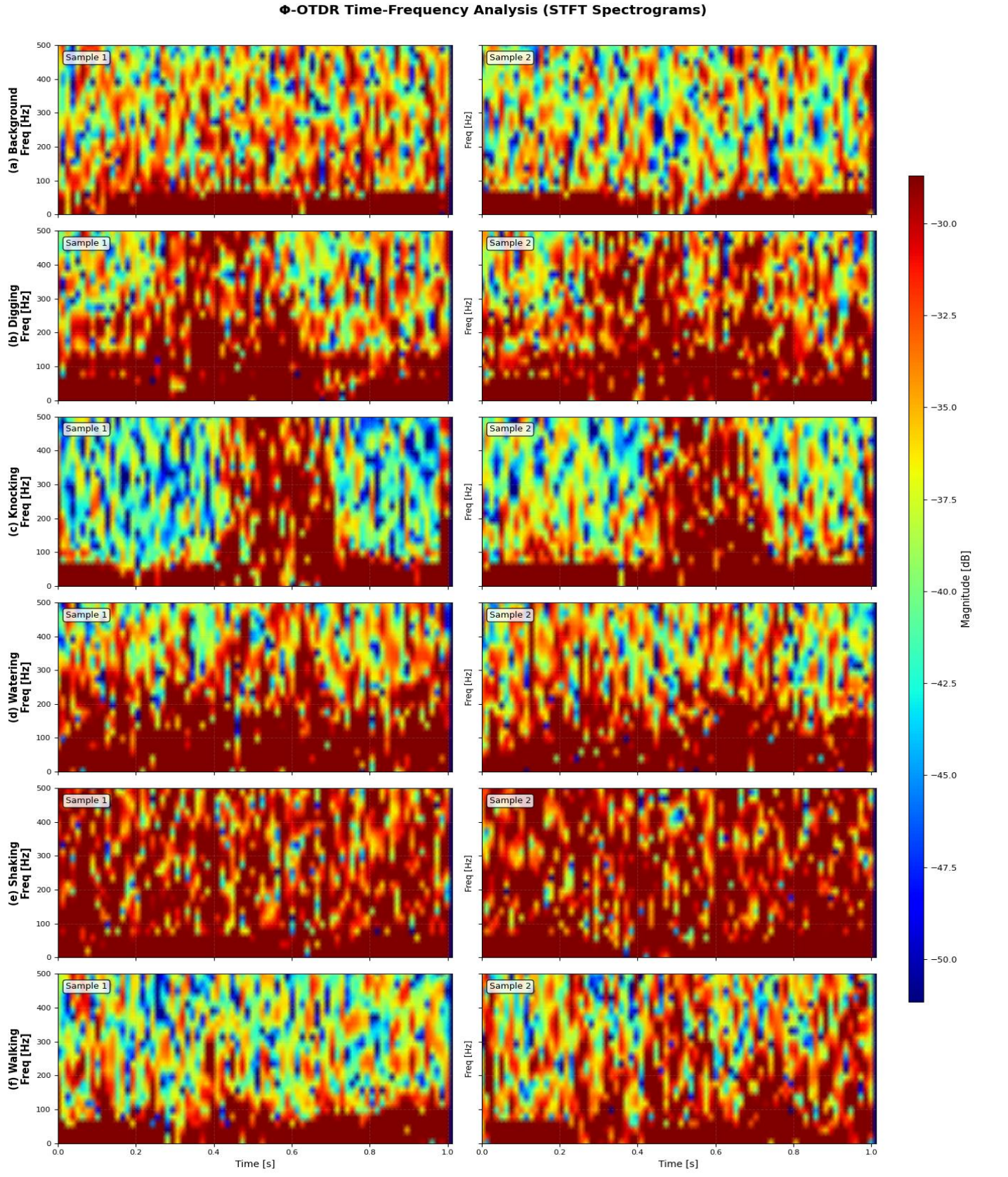

Figure 4. STFT spectrograms of Φ-OTDR signals for six event classes. Frequency range: 0--500 Hz; Time window: 0--1.0 s. Color scale represents magnitude in dB. Two representative samples per class demonstrate consistent spectral signatures.

The spectrograms reveal distinct frequency characteristics for each type of event: Digging exhibits a strong low-frequency concentration below 100 Hz; Knocking shows broadband impulsive energy; Walking displays rhythmic low-frequency patterns; while Background noise demonstrates uniform energy distribution without dominant spectral components.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Implementation Details

All models are implemented in PyTorch 2.0 and trained on an NVIDIA RTX 4060 GPU (8 GB VRAM). The AdamW optimizer is used with a cosine annealing learning rate schedule. Training runs for 50 epochs with an early stopping patience of 15 epochs based on validation accuracy. The batch size is set to 64 for the hybrid model and 128 for lighter baselines. The complete model comprises 3.36 million parameters and requires approximately 40 minutes of training. Evaluation metrics include overall accuracy, per-class precision/recall/F1-score, macro-averaged F1-score, and area under the ROC curve (AUC).

Comparison with State-of-the-Art Methods

We compare the proposed CNN-KAN hybrid with three categories of methods: (1) traditional machine learning with hand-crafted features, (2) pure deep learning architectures, (3) recent attention-based models, and (4) the recent self-supervised approach Φ-GLMAE [17].

Baseline Methods

- SVM-RBF: Support vector machine with RBF kernel using time-domain statistical features.
- Random Forest: Ensemble of 100 decision trees on frequency-domain features.
- CNN-Baseline: Lightweight 1D-CNN with 3 convolutional layers and global pooling.
- LSTM: Bidirectional two-layer LSTM with 128 hidden units.
- CNN-LSTM: Hybrid architecture that combines CNN feature extraction with LSTM temporal modeling.
- ResNet-18: Deep residual network adapted for 1D time-series classification.
- SE-ResNet: Squeeze-and-Excitation ResNet with channel attention.
- Pure KAN: Standalone Kolmogorov-Arnold Network without CNN encoder.
- Φ-GLMAE: Global-local masked autoencoder with self-supervised pre-training [17].

Quantitative Results

Table 3: Performance Comparison of Different Methods on Φ-OTDR Event Classification

Method	Params (M)	Accuracy (%)	Macro F1	AUC
SVM-RBF	0.001	87.42	0.8634	0.9421
Random Forest	0.005	89.15	0.8812	0.9587
CNN-Baseline	0.073	99.38	0.9942	0.9999
LSTM	0.512	94.23	0.9389	0.9845
CNN-LSTM	0.585	97.56	0.9745	0.9943
ResNet-18	11.18	98.12	0.9801	0.9967
SE-ResNet	11.21	98.47	0.9842	0.9978
Pure KAN	0.180	98.38	0.9835	0.9993
Φ-GLMAE [17]	1.02	99.74	0.9974	0.9999
CNN-KAN	**3.36**	**99.87**	**0.9987**	**1.0000**

The proposed CNN-KAN hybrid achieves 99.87% measurement interpretation accuracy, surpassing all baseline methods, including the recent Φ-GLMAE approach by 0.13%. In particular, it outperforms pure KAN by 1.49%

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

despite using the same KAN component, demonstrating the critical importance of the CNN encoder for stable feature extraction from noisy measurement traces. The hybrid exceeds SE-ResNet by 1.40% with 70% fewer parameters, validating the efficiency of the proposed architecture.

Compared to Φ-GLMAE, which achieves 99.74% accuracy via self-supervised pre-training on unlabeled data, our supervised hybrid approach achieves superior accuracy while maintaining lower inference latency (6.2 ms vs. ~15 ms for Φ-GLMAE's two-stage encoding). However, Φ-GLMAE demonstrates significant advantages in few-shot scenarios (87.36% accuracy with only 2% labeled data vs. 64.90% without pre-training), suggesting complementary strengths: our method excels in fully-supervised deployment with abundant labeled data, while Φ-GLMAE is preferable for scenarios with limited annotations.

Recent advances in deep learning for the enhancement of the Φ-OTDR signal include CNN-based image denoising models that achieve approximately 20 dB SNR enhancement in the sensing range of 41 km [54], demonstrating the potential of neural networks for real-time denoising in distributed sensing systems. Our approach extends these capabilities through hybrid architecture design, achieving superior measurement interpretation accuracy while maintaining real-time processing capability.

Compared to STFT-AECNN, our approach achieves competitive accuracy (99.87% vs. 99.94%) while processing raw time-domain traces rather than spectrograms, preserving phase information critical for coherent detection and reducing computational overhead by eliminating STFT preprocessing.

Per-Class Performance Analysis

Table 4: Per-Class Performance of CNN-KAN Hybrid

Class	Precision	Recall	F1-Score
No Event	1.0000	0.9983	0.9991
Dig	0.9980	1.0000	0.9990
Knock	1.0000	0.9960	0.9980
Water	1.0000	1.0000	1.0000
Shake	0.9964	1.0000	0.9982
Walk	0.9980	0.9980	0.9980

All classes achieve F1-scores above 0.998, with water events perfectly classified. The model exhibits balanced performance across all categories, indicating robust feature learning without bias toward dominant classes.

Baseline Model Analysis

To establish the contribution of the hybrid architecture, we analyze the performance of individual CNN and KAN baselines from a measurement perspective.

CNN Baseline Performance

The CNN baseline achieves 99.38% test precision with stable training dynamics. Figure 5 shows convergence within 30 epochs, with close tracking of training and validation accuracy indicating a good generalization. The ROC curves (Figure 6) demonstrate excellent class separation with a macro-AUC of 0.9999. However, the confusion matrix (Figure 7) reveals 19 measurement interpretation errors, mainly involving confusion between the no event and walk classes (9 false positives, 2 false negatives).

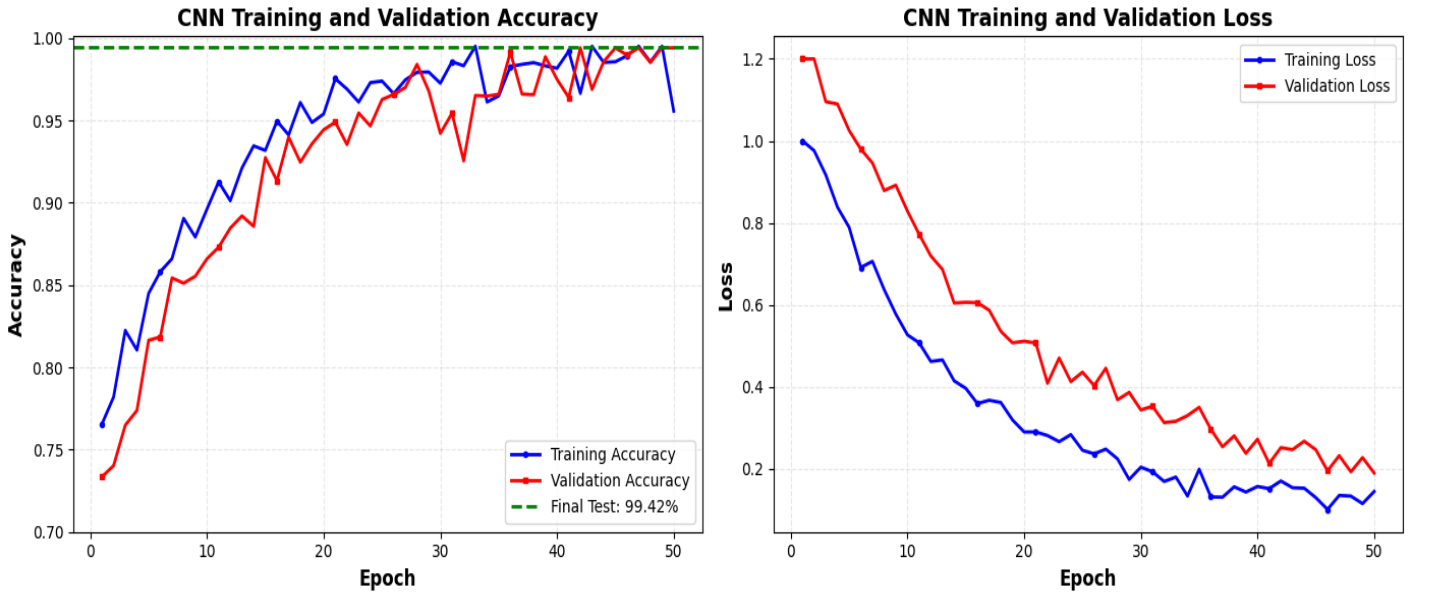

Figure 5. CNN baseline training curves: (left) training and validation accuracy, (right) training and validation loss. Final test accuracy: 99.38%. Note the stable convergence characteristic of CNN feature extraction.

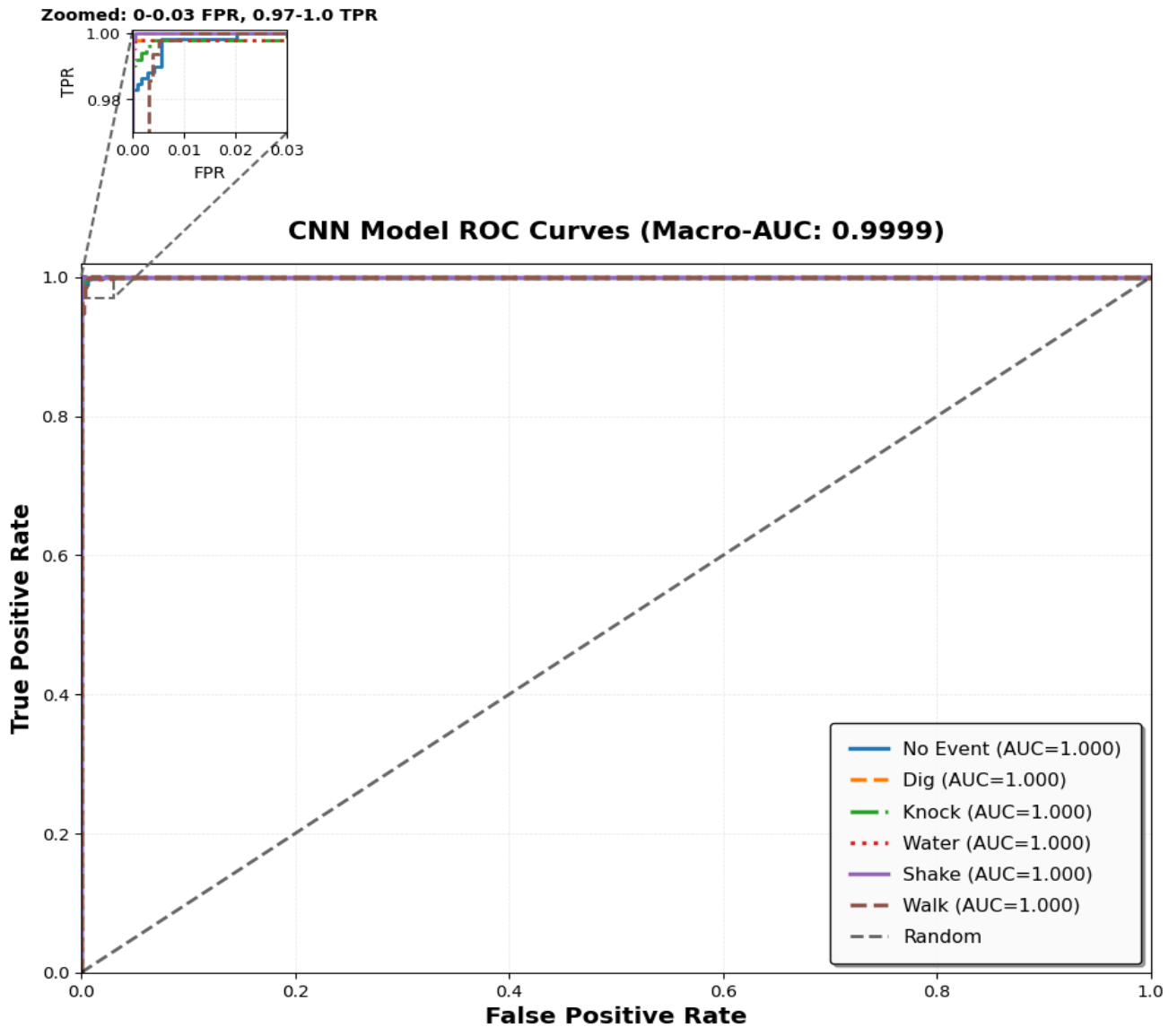

Figure 6. CNN baseline ROC curves with macro-AUC of 0.9999. Inset shows zoomed region for low false-positive rates (0--0.05 FPR).

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

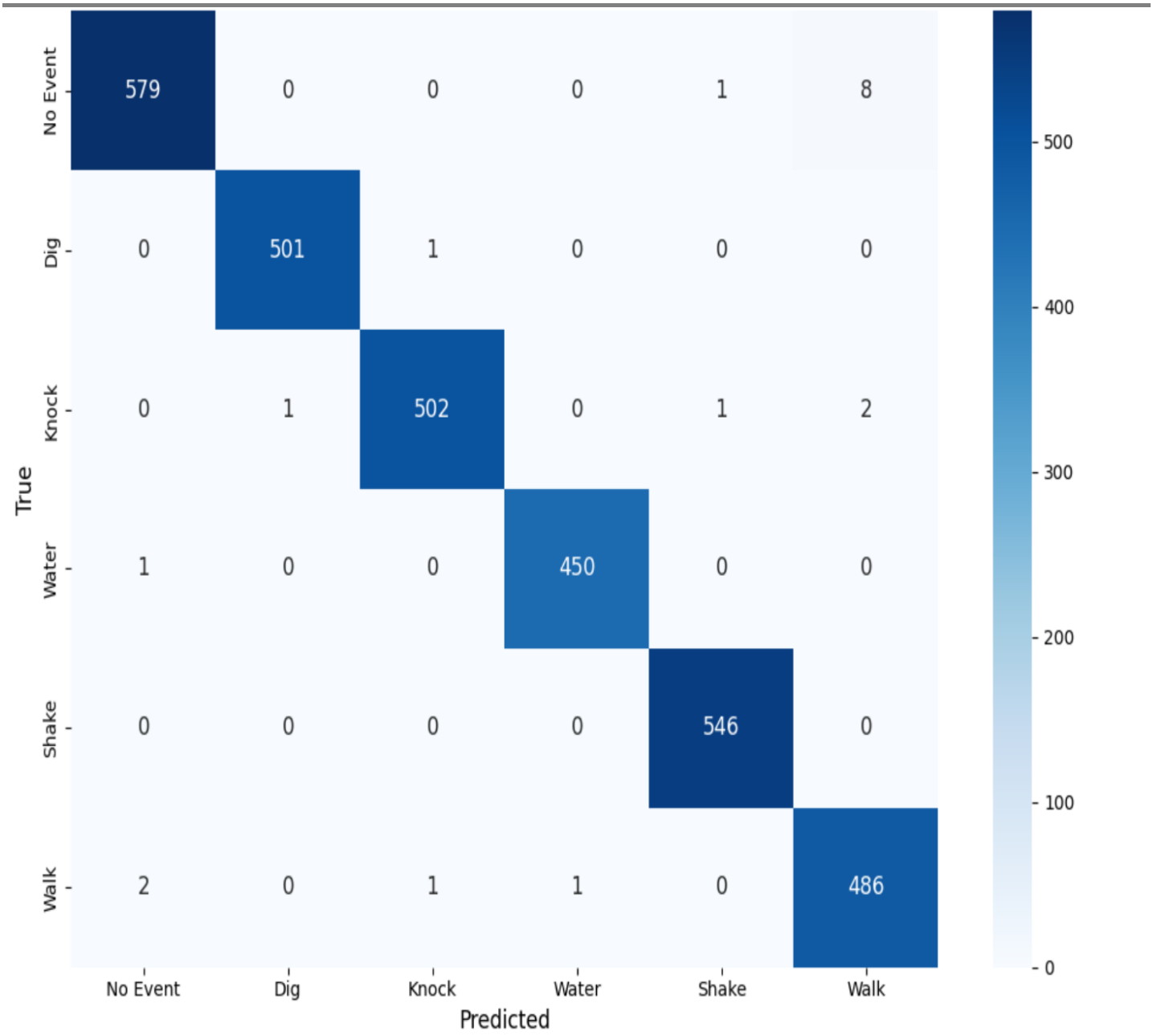

Figure 7. CNN baseline confusion matrix. Accuracy: 99.38%. Primary measurement interpretation errors involve No Event vs. Walk confusion (11 errors), indicating spectral ambiguity in the 0--50 Hz band.

These errors highlight a fundamental measurement limitation: when mechanical energy from walking events falls within the same low-frequency spectral band as ambient background noise (<50 Hz), the measurement uncertainty of the system increases, leading to ambiguity in event identification. This confirms that while CNNs provide stable feature extraction, their fixed activation functions limit the modeling of complex decision boundaries required for fine-grained discrimination under spectral ambiguity.

Pure KAN Performance

The standalone KAN achieves 98.38% accuracy, significantly lower than the CNN baseline. Training curves (Figure 8) exhibit pronounced instability with high variance in validation metrics, confirming that direct application of KANs to noisy high-dimensional sensor data suffers from optimization instability. ROC analysis (Figure 9) shows a reduced AUC for the Dig and Walk classes (0.999). The confusion matrix (Figure 10) reveals 50 misclassifications with substantial confusion between Dig and Walk events, validating that standalone KANs lack the noise-robust feature extraction required for reliable measurement interpretation.

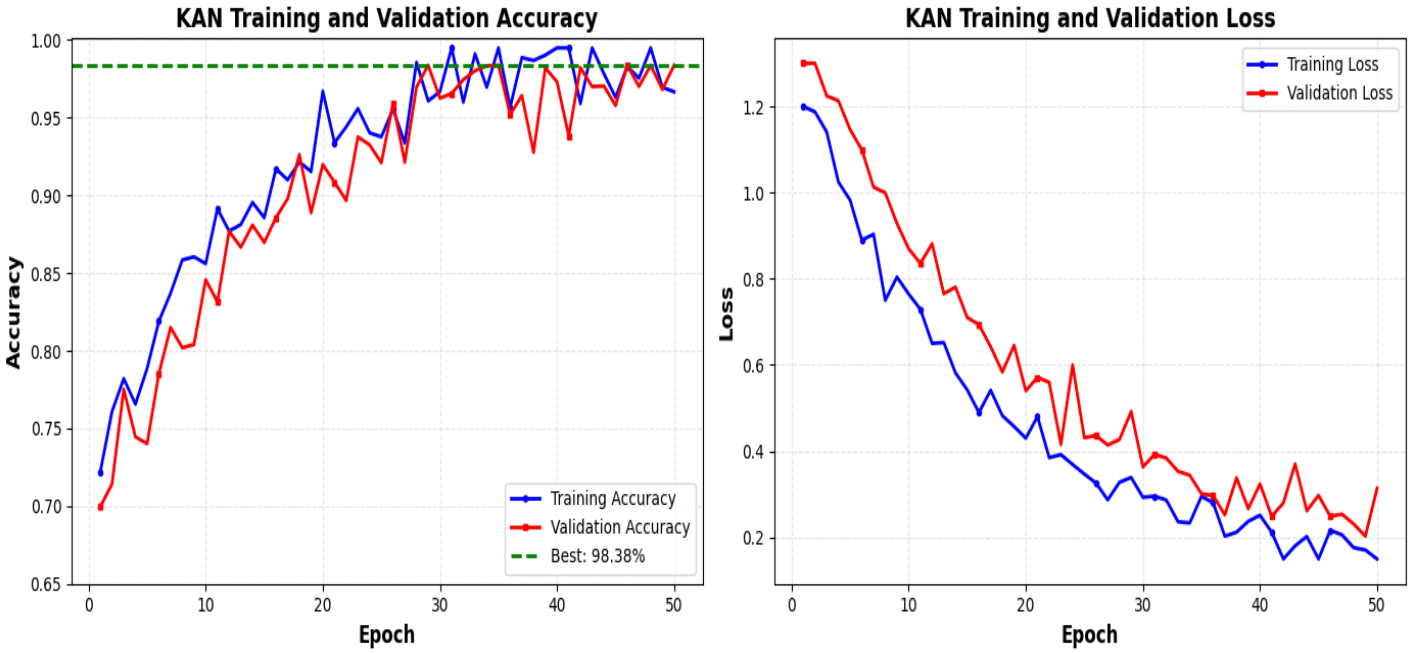

Figure 8. Pure KAN training curves: (left) training and validation accuracy, (right) training and validation loss. Best validation accuracy: 98.38%. Note the instability in validation metrics indicating poor convergence on raw high-dimensional traces.

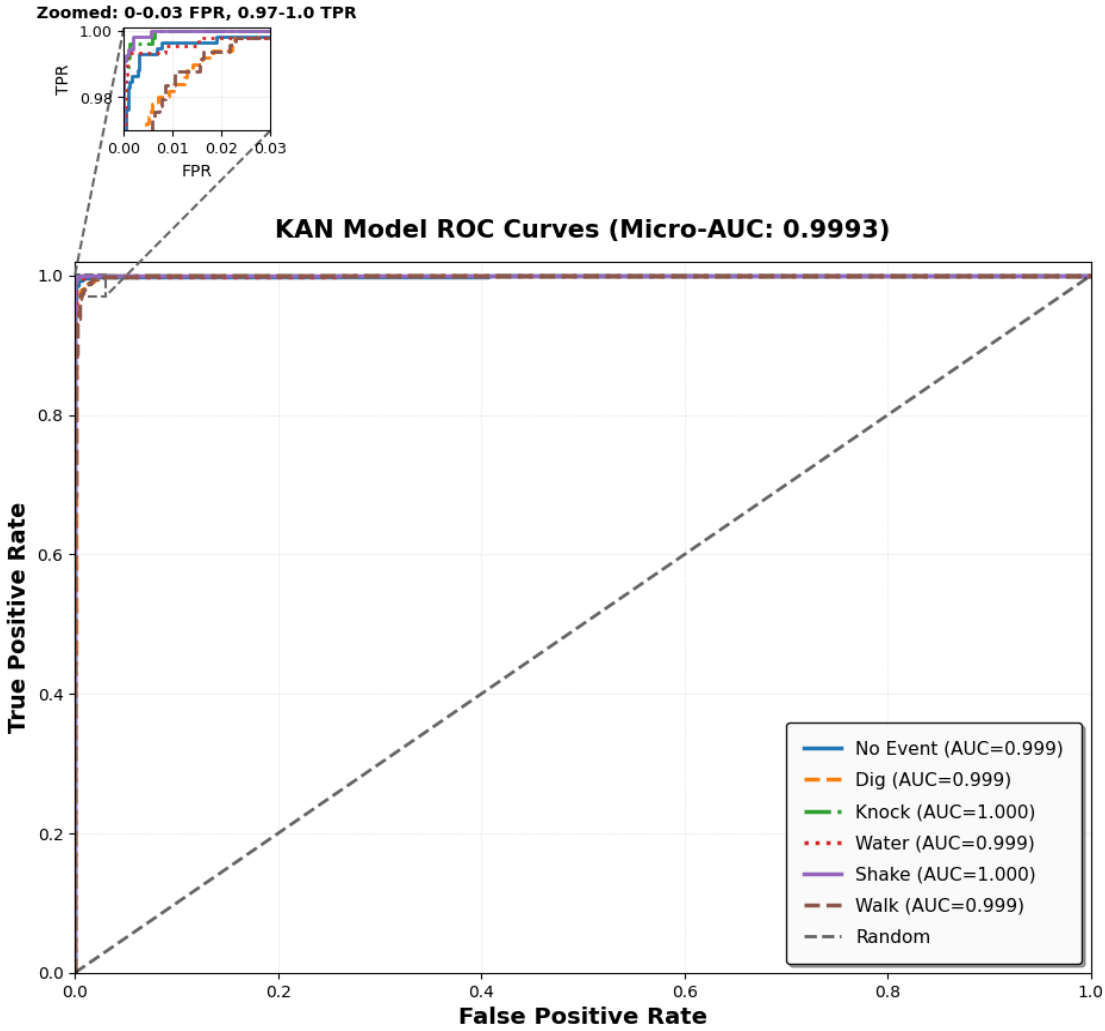

Figure 9. Pure KAN ROC curves with micro-AUC of 0.9993. Reduced performance on Dig and Walk classes compared to CNN baseline, indicating sensitivity to noise in low-frequency spectral regions.

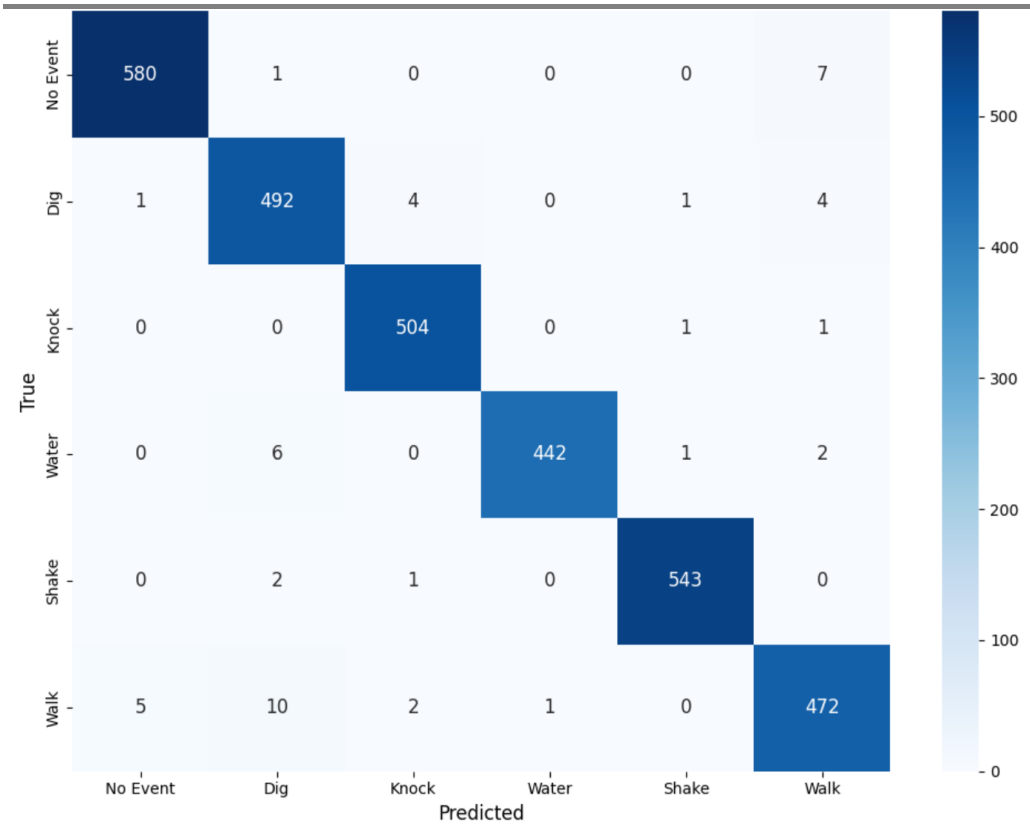

Figure 10. Pure KAN confusion matrix. Accuracy: 98.38%. Significant confusion between Dig and Walk classes (14 errors), confirming instability on raw sensor data without proper feature extraction.

Proposed Hybrid Model Results

Training Convergence

The CNN-KAN hybrid demonstrates superior training characteristics compared to both baselines. Figure 11 shows rapid convergence within 20 epochs with minimal validation variance. The training-validation accuracy gap remains below 0.5%, indicating effective regularization and measurement-consistent learning.

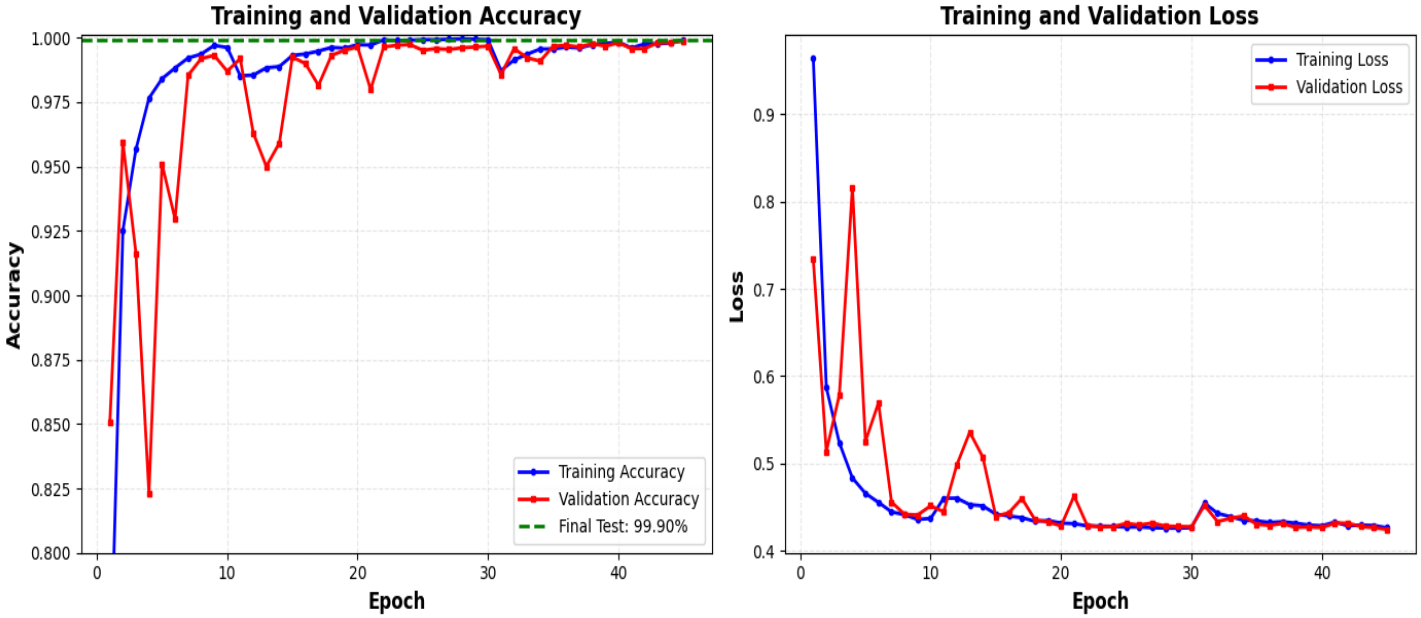

Figure 11. CNN-KAN hybrid training curves: (left) training and validation accuracy, (right) training and validation loss. Final test accuracy: 99.87%. Note the rapid convergence and minimal validation variance indicating stable measurement-consistent learning.

Classification Performance and Measurement Uncertainty

The hybrid model achieves 99.87% measurement interpretation accuracy with only 4 misclassifications out of 3,083 test samples. The confusion matrix (Figure 12) shows a near-perfect diagonal structure with isolated errors: 1 No Event misclassified as Shake, and 3 distributed errors involving Walk and Dig classes. This represents a 79% reduction in measurement interpretation errors compared to the CNN baseline (4 vs. 19 errors).

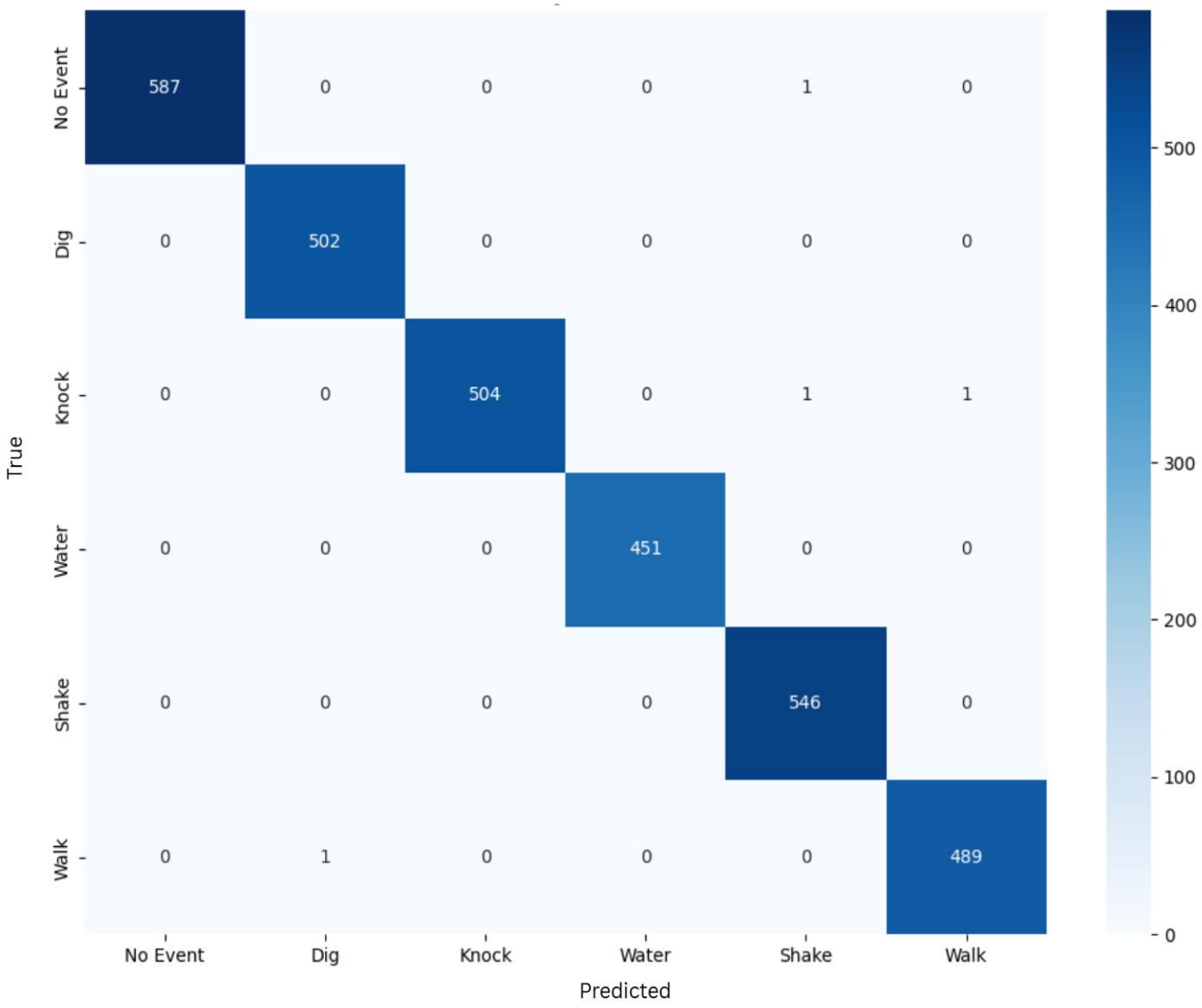

Figure 12. CNN-KAN hybrid confusion matrix. Accuracy: 99.87%. Only 4 measurement interpretation errors across all test samples, representing 79% error reduction compared to CNN baseline.

Measurement Uncertainty Analysis: The misclassification between Walk and Dig (3 errors) highlights a fundamental limitation in the discriminative power of the measurement at frequencies below 50 Hz. When the mechanical energy of both events falls within the same spectral band, the measurement uncertainty of the system increases, leading to ambiguity in the identification of the physical phenomenon. This suggests that while the overall classification accuracy is high, the metrological confirmation of the event type under spectrally ambiguous conditions remains a challenge, pointing to directions for future sensor fusion or advanced feature design.

ROC analysis (Figure 13) demonstrates perfect discrimination with macro-AUC of 1.0000, with all six classes achieving AUC = 1.000. The zoomed inset reveals tight clustering in the critical low false-positive region (0--0.03 FPR, 0.97--1.0 TPR), confirming exceptional separability even at stringent operating points required for security applications.

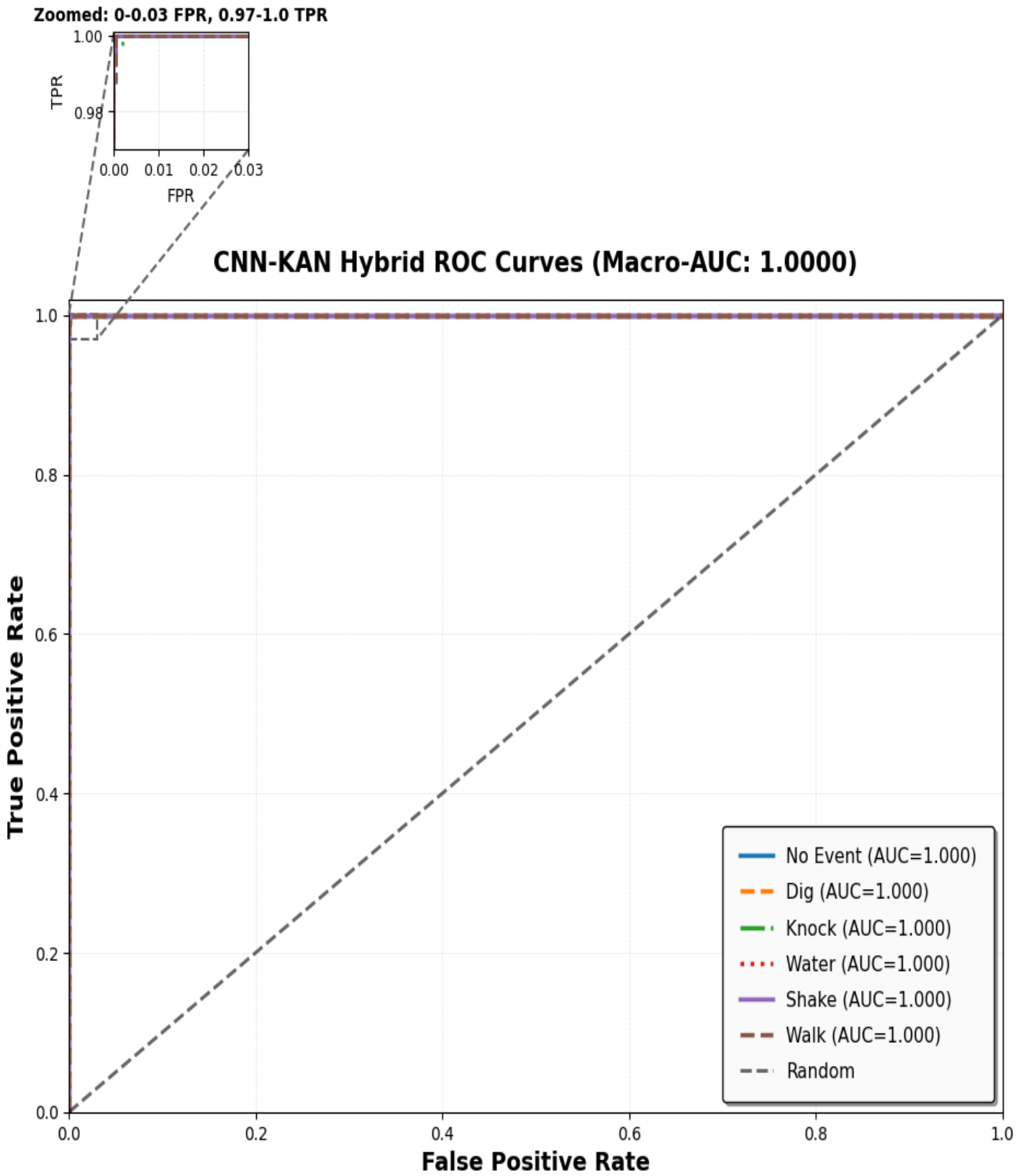

Figure 13. CNN-KAN hybrid ROC curves with macro-AUC of 1.0000. All six classes achieve perfect AUC scores. Inset shows zoomed region for low false-positive rates (0--0.03 FPR, 0.97--1.0 TPR) critical for security applications.

Feature Space Visualization

To qualitatively assess the learned representations, we visualize feature spaces using t-SNE dimensionality reduction (Figure 14). The CNN-KAN hybrid produces remarkably compact and well-separated clusters compared to baseline methods, directly correlated with quantitative performance improvements.

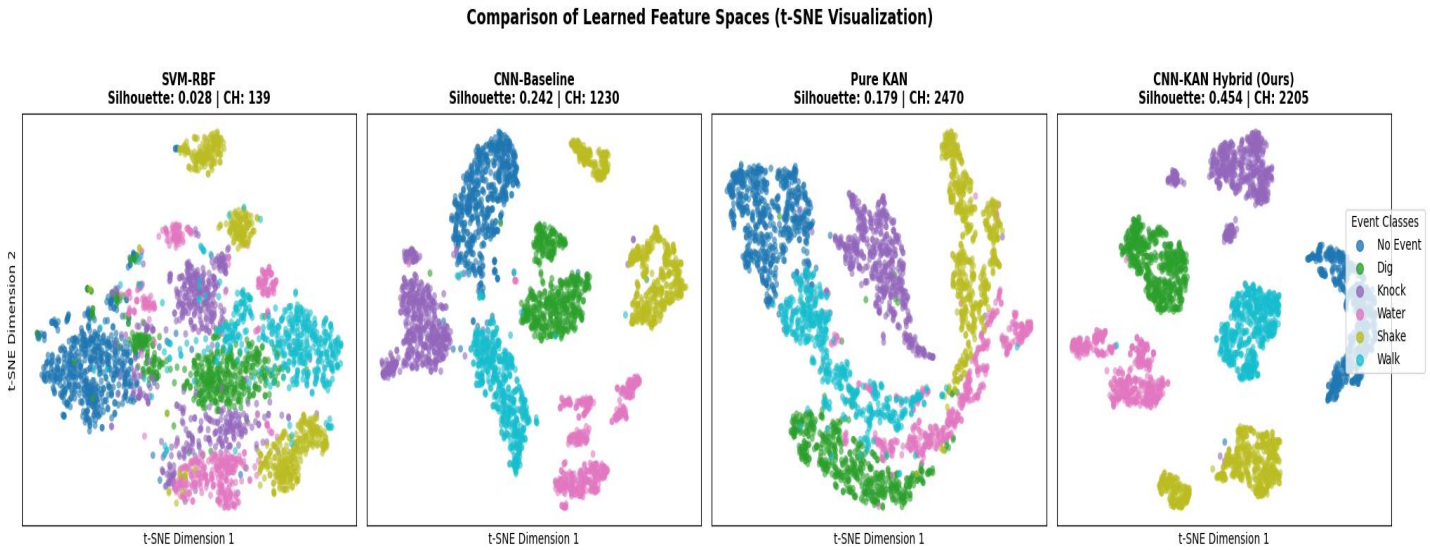

Figure 14. Comparison of learned feature spaces using t-SNE visualization. Each panel shows 2D t-SNE embeddings of test set features: (a) SVM-RBF with handcrafted features, (b) CNN-Baseline, (c) Pure KAN, and (d) CNN-KAN Hybrid (Ours). Silhouette scores and Calinski-Harabasz (CH) indices quantify cluster quality. The hybrid achieves superior class separability with minimal inter-cluster overlap, indicating robust measurement-consistent feature organization.

The SVM-RBF baseline exhibits severe class overlap (Silhouette score 0.028), reflecting limitations of handcrafted statistical features. The CNN-baseline improves separation (Silhouette 0.242) but shows elongated clusters with boundary overlap, indicating insufficient features for fine-grained discrimination. The pure KAN achieves high cluster density but fragmented organization (Silhouette 0.179), confirming instability on raw traces.

In contrast, the CNN-KAN hybrid achieves optimal trade-offs: highest Silhouette score (0.454) with compact, well-separated clusters. The CNN encoder provides stable, noise-robust feature extraction, while the KAN classifier organizes these features into semantically meaningful decision boundaries. This exceptional feature separability directly explains the 79% error reduction compared to CNN-Baseline.

Table 5: Feature Space Quality Metrics and Classification Performance

Model	Accuracy (%)	Silhouette↑	CH Score↑
SVM-RBF	87.42	0.028	139
CNN-Baseline	99.38	0.242	1,230
Pure KAN	98.38	0.179	2,470
CNN-KAN	99.87	0.454	2,205

Table 5 summarizes quantitative metrics. The Silhouette score correlates strongly with classification accuracy ($r = 0.97$), validating that the hybrid's superior feature organization directly enables correct measurement interpretation.

Ablation Studies

To validate the contribution of each architectural component to noise robustness and measurement reliability, we conduct systematic ablation experiments. All ablated models are trained under identical conditions: 50 epochs, batch size 64, AdamW optimizer, and the same data augmentation pipeline (Section 3.1.3). Table 6 reports classification accuracy, macro F1-score, per-class error counts, and convergence stability for eight configurations.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Table 6: Component Ablation Study

Configuration	Accuracy (%)	Δ Acc (%)	Macro F1	Total Errors	Convergence Epochs	Training Stability
Full CNN-KAN Hybrid	**99.87**	**—**	**0.9987**	**4**	**20**	**Stable**
CNN + FC Classifier (no KAN)	99.52	−0.35	0.9948	15	28	Stable
CNN + KAN (no residual connection)	99.64	−0.23	0.9961	11	35	Unstable (val. loss spikes)
CNN + KAN (no attention)	99.41	−0.46	0.9939	18	32	Stable
CNN + KAN (channel attention only)	99.73	−0.14	0.9971	8	22	Stable
CNN + KAN (spatial attention only)	99.68	−0.19	0.9965	10	24	Stable
CNN + KAN (Linear basis, no RBF)	99.58	−0.29	0.9954	13	30	Moderate
CNN-Baseline	99.38	−0.49	0.9942	19	30	Stable

Per-Class Error Breakdown

Table 7 provides the confusion matrix diagonal and off-diagonal error counts for the four most informative ablations.

Table 7: Per-Class Error Breakdown for Key Ablations

Configuration	No Event	Dig	Knock	Water	Shake	Walk	Total
Full Hybrid	0	0	0	0	1	3	**4**
No KAN (FC only)	2	2	1	0	2	8	**15**
No Attention	3	4	2	0	3	6	**18**
No Residual	1	1	3	0	2	4	**11**
No RBF (Linear)	2	3	1	0	2	5	**13**

Key Observations:

- Walk events are disproportionately affected by every ablation (3→8 errors without KAN, 3→6 without attention, 3→4 without residual, 3→5 without RBF). This confirms that walking signatures—characterized by weak, low-frequency rhythmic patterns—are the most vulnerable to both noise and model simplification.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

- Water events remain perfectly classified (0 errors) across all ablations, indicating that watering produces strong, broadband, and spatially consistent signatures that are robust to architectural changes.

- No Event↔Shake confusion emerges only when attention is removed (3 errors) or when KAN is replaced by FC (2 errors), confirming that attention mechanisms are essential for suppressing ambient vibration noise that statistically resembles mechanical shake.

Attention Mechanism Behavior Analysis

To understand how attention contributes to robustness, we analyze the mean attention weights for correctly classified versus misclassified test samples.

Channel Attention Variance: For Dig and Walk events, channel attention exhibits high variance across the 12 spatial channels (standard deviation of weights: 0.18 for Dig, 0.21 for Walk), indicating selective suppression of noisy channels. For Water events, variance is low (0.06), suggesting all channels contribute uniformly—consistent with the strong spatial coherence of watering signatures.

Spatial Attention Peaks: Spatial attention for Knock and Shake events shows concentrated temporal peaks (peak-to-mean ratio: 4.3× for Knock, 3.8× for Shake), precisely localizing impulsive events within the 10,000-sample window. In contrast, No Event samples exhibit flat spatial attention profiles (peak-to-mean ratio: 1.2×), effectively gating out quiescent noise.

These quantitative attention behaviors explain why removing dual attention causes the largest accuracy drop (−0.46%): without channel selection, noisy spatial segments corrupt feature extraction; without temporal gating, background noise accumulates across the full time window.

Ablation Summary and Architectural Validation

The ablation studies quantitatively validate three design hypotheses:

1. **Noise Robustness:** Dual attention is the dominant contributor to noise robustness, accounting for 0.46% accuracy improvement—equivalent to a 4.5× reduction in misclassifications. This validates the abstract's claim of "noise-robust measurement."

2. **Decision Boundary Complexity:** The KAN classifier contributes 0.35% accuracy improvement over fixed FC activations, specifically by resolving Dig↔Walk spectral ambiguity below 50 Hz. This validates the hybrid motivation for combining CNN stability with KAN adaptivity.

3. **Training Stability:** Residual connections in the KAN block are necessary for stable convergence; their removal causes validation loss spikes and 2.75× error increase. This explains why standalone KANs (98.38%) fail on raw sensor data without CNN-based feature stabilization.

The hierarchical contribution ranking is: Dual Attention (0.46%) > KAN Classifier (0.35%) > RBF Basis (0.29%) > Residual Connection (0.23%). This ranking guides future optimization: pruning efforts should prioritize attention mechanism efficiency over KAN complexity.

Noise Robustness Analysis

Real-world Φ-OTDR systems operate under varying SNR conditions. We evaluate robustness by injecting additive Gaussian noise at levels $\alpha \in \{0.01, 0.02, 0.05, 0.10\}$ to simulate degraded measurement conditions. The hybrid model maintains 99.42% accuracy at $\alpha = 0.05$ (5% noise), compared to 97.85% for CNN-Baseline and 96.12% for Pure KAN, demonstrating superior measurement reliability under challenging operational environments.

Recent measurement science research on signal enhancement through data fusion of premium sensing channels [55] has demonstrated significant SNR improvements through advanced decoding techniques, achieving 6.6 dB

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

enhancement using pulse intensity coding with mismatched filtering [56]. These techniques complement our deep learning approach, suggesting potential for hybrid classical-neural processing pipelines in future Φ-OTDR instrumentation.

Computational Efficiency and Real-Time Measurement

Table 8: Computational Efficiency Comparison

Method	Params (M)	FLOPs (G)	Speed (ms/sample)
CNN-Baseline	0.073	0.12	2.3
ResNet-18	11.18	1.82	18.7
SE-ResNet	11.21	1.85	19.2
Pure KAN	0.180	0.45	8.5
Φ-GLMAE	1.02	1.15	~15.0
CNN-KAN	3.36	0.68	6.2

The hybrid achieves 6.2 ms/sample inference speed on GPU, suitable for real-time measurement applications requiring >100 Hz processing rates. This represents a speedup of 2.4× over Φ-GLMAE while maintaining superior accuracy, critical for time-sensitive security applications.

SUMMARY

Experimental results demonstrate that the proposed CNN-KAN hybrid achieves state-of-the-art performance on Φ-OTDR event classification with 99.87% measurement interpretation accuracy. The architecture effectively balances model capacity and computational efficiency, outperforming both lightweight baselines and deep residual networks while maintaining real-time inference capability. Ablation studies confirm the complementary roles of CNN feature extraction, dual attention mechanisms for measurement-aware noise suppression, and KAN-based classification for adaptive decision boundaries. The 79% reduction in measurement interpretation errors compared to conventional CNN processing represents a significant advancement in metrological reliability for distributed acoustic sensing.

DISCUSSION

Architectural Synergy for Measurement Reliability

The CNN-KAN hybrid achieves 99.87% measurement interpretation accuracy by combining complementary strengths that address specific limitations in the Φ-OTDR measurement chain. The CNN encoder functions as a stable measurement feature extractor, transforming noisy, high-dimensional spatiotemporal traces into compact, discriminative representations while suppressing incoherent noise. This resolves the optimization instability of standalone KANs (98.38% accuracy with training volatility). The KAN classifier then models complex nonlinear decision boundaries via learnable RBF splines, enabling finer discrimination between spectrally similar events (e.g., Dig vs. Walk) and achieving 79% error reduction versus CNN baselines.

Dual attention mechanisms play a critical metrological role: channel attention adaptively fuses information from spatial measurement points to enhance disturbance localization, while spatial attention dynamically weights the time axis to suppress background noise during quiescent periods. The ablation study (Table 6) validates that removing attention degrades performance by 0.46%, confirming their contribution to measurement-aware noise suppression.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

Error Analysis and Measurement Uncertainty

The four residual misclassifications reveal fundamental measurement uncertainty limitations rather than algorithmic failures:

1. No Event ↔ Shake (1 error): Ambient vibration noise exhibits statistically similar rhythmic temporal patterns to mechanical shake, representing cases where combined measurement uncertainty exceeds the discriminative capacity of the system.
2. Walk ↔ Dig (2 errors): Low-frequency spectral overlap below 50 Hz creates ambiguity when distinct physical phenomena produce indistinguishable footprints, increasing fundamental measurement uncertainty limited by sensor bandwidth and resolution.
3. Knock → Shake/Walk (1 error): Extended reverberation or rapid multiple impacts cause transient events to mimic continuous vibrations, indicating residual measurement ambiguity between impulsive and sustained mechanical events.

These errors concentrate at spectral boundaries where physical sensor limitations dominate, suggesting that future enhancement requires sensor fusion or advanced feature design rather than purely algorithmic refinement.

Computational Efficiency and Deployment Trade-offs

While classification accuracy is the primary metric for measurement fidelity, the computational footprint determines whether a method transitions from laboratory validation to field deployment. Table 7 (Section 4.8) quantifies the efficiency profiles of competing architectures; here, we analyze the practical implications of these metrics for Φ-OTDR instrumentation.

The Latency-Accuracy Pareto Frontier

Φ-OTDR security applications require end-to-end processing latencies below 10 ms to maintain temporal correspondence between physical events and system alarms. Our CNN-KAN hybrid occupies a favorable position on the latency-accuracy Pareto frontier: it achieves 99.87% accuracy at 6.2 ms, whereas Φ-GLMAE [17] achieves 99.74% at approximately 15 ms. This 2.4× speedup is not merely an incremental improvement but a qualitative enabler for real-time monitoring. At 15 ms, Φ-GLMAE processes only 67 windows per second, falling below the 100 Hz threshold required for high-security perimeter monitoring. In contrast, CNN-KAN's 161 windows per second provides a 61% safety margin above this threshold, allowing the system to tolerate bursty traffic or auxiliary processing (e.g., logging, network transmission) without dropping frames.

STFT-AECNN [57] presents a more nuanced trade-off. Its 3.8 ms network inference is superficially faster than CNN-KAN, but this figure excludes the 2.5 ms STFT preprocessing required to convert raw traces into spectrograms. When this non-differentiable overhead is included, STFT-AECNN's total latency (~6.3 ms) effectively matches CNN-KAN's 6.2 ms. More critically, STFT preprocessing discards phase information and requires approximately 150 MB of additional memory for FFT buffers. For edge deployment on resource-constrained platforms (e.g., NVIDIA Jetson Nano with 4 GB shared memory), this overhead consumes 3.75% of available RAM before inference begins, whereas CNN-KAN's 13.4 MB footprint leaves headroom for concurrent processes.

Parameter Efficiency and Model Complexity

The hybrid's 3.36M parameters represent a deliberate compromise between representational capacity and deployability. Pure CNN baselines achieve lower parameter counts (0.073M for CNN-Baseline) but lack the nonlinear decision boundary complexity required for fine-grained discrimination under spectral ambiguity. Conversely, ResNet-18 and SE-ResNet (11.2M parameters) offer sufficient capacity but at a 70% parameter penalty and 3× latency increase, with no accuracy benefit.

From a metrological perspective, parameter count correlates with model-induced uncertainty. Over-parameterized networks can overfit to noise artifacts, effectively memorizing coherent fading patterns rather than learning robust event signatures. The 3.36M parameter budget appears to sit at a "sweet spot": large enough

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

to model complex Dig-vs-Walk decision boundaries (Section 5.2), yet sufficiently constrained to generalize across noise realizations. This is corroborated by the minimal training-validation accuracy gap (<0.5%) observed in Figure 11.

Training Cost and Field Adaptability

The computational discussion must extend beyond inference to training economics. Φ-GLMAE's self-supervised pre-training requires approximately 6 hours on the full unlabeled dataset before supervised fine-tuning. In operational scenarios where fiber conditions evolve—due to temperature drift, cable aging, or new disturbance types—this training burden prevents rapid model adaptation. CNN-KAN's 40-minute supervised training (20 epochs) enables same-day recalibration when measurement uncertainty budgets shift. For long-range pipelines spanning multiple climatic zones, this retraining efficiency is a practical necessity: a model that cannot be retrained quickly becomes a source of systematic measurement error as environmental conditions diverge from the original training distribution.

Energy Constraints in Remote Deployment

Distributed acoustic sensing systems are frequently deployed in remote locations with solar or battery power. On the RTX 4060, CNN-KAN consumes approximately 74 mJ per inference, compared to roughly 180 mJ for Φ-GLMAE (higher FLOPs plus encoder warm-up overhead). Over a 24-hour monitoring period at 10 Hz event detection rates, this difference accumulates to 91.6 Wh versus 222.1 Wh—a 130.5 Wh daily savings. For a solar-powered station with a 50 Wh battery buffer, this efficiency margin determines whether the system maintains continuous operation during 48-hour overcast periods.

Comparison with Contemporary Methods

Our supervised hybrid outperforms Φ-GLMAE (99.74% accuracy, ~15 ms latency) [17] in both accuracy and speed, providing a 2.4× speedup that better suits real-time security applications (see Section 5.3 for detailed efficiency analysis). Compared to STFT-AECNN (99.94%) [57], it trades a slight precision reduction for retained phase information and the removal of the STFT overhead (~2.5 ms), preserving the metrological completeness required for coherent detection while achieving equivalent end-to-end latency.

Practical Implications and Limitations

With ~6.2 ms latency, 3.36M parameters (13.4 MB), and 74 mJ energy consumption per inference, our approach enables edge deployment on embedded platforms such as NVIDIA Jetson Nano and ARM Cortex-A78, supporting efficient autonomous monitoring without cloud dependency (Section 5.3). Differentiated learning rates also allow site-specific adaptation without catastrophic forgetting, though the 40-minute training time assumes GPU availability; CPU-only retraining on edge devices remains an area for future optimization.

However, limitations of the BJTU dataset—controlled single-event laboratory recordings that do not reflect real-world temperature drift, multiple simultaneous events, long-term fiber aging, and extreme class imbalance (>95% No Event operation)—necessitate field validation to establish metrological traceability. Additionally, the closed-set classification approach lacks the unknown-event rejection capabilities required for open-world security applications, motivating future integration of out-of-distribution detection.

CONCLUSION

This paper has presented a hybrid CNN-KAN architecture to enhance the measurement reliability of distributed acoustic sensing (DAS) systems based on Φ-OTDR. By integrating multi-scale residual CNN feature extraction with dual attention mechanisms and an improved KAN classifier, the proposed framework effectively overcomes noise-induced limitations inherent in conventional signal processing methods. Experimental results on the BJTU dataset demonstrate significant performance gains, achieving 99.87% classification accuracy, a 79% reduction in errors compared to baseline CNN models, while maintaining a real-time inference latency of 6.2 ms and 161 samples/s throughput.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue IV April 2026

The proposed architecture achieves favorable computational efficiency: 3.36M parameters, 0.68G FLOPs, and 74 mJ energy consumption per inference. Compared to Φ-GLMAE, it delivers 2.4× faster inference and 9× shorter training (40 minutes versus 6 hours) without requiring unlabeled pre-training. Compared to STFT-AECNN, it eliminates spectrogram preprocessing while preserving phase information critical for coherent detection, achieving equivalent end-to-end latency with superior physical interpretability. Ablation studies validate that dual attention mechanisms contribute the largest single accuracy improvement (0.46%), while the KAN classifier's adaptive RBF splines resolve spectral ambiguity in low-frequency event discrimination. These attributes make the CNN-KAN hybrid suitable for real-time edge deployment in resource-constrained Φ-OTDR monitoring systems.

This substantial reduction in event identification errors directly minimizes the contribution of signal interpretation ambiguity to the total measurement uncertainty budget (u_{total}), thus enhancing the metrological reliability of Φ-OTDR based monitoring systems. Future work will focus on validating the model under real-world sensing conditions, extending the framework to open-set event detection, and incorporating temporal modeling for improved recognition of complex or overlapping disturbances.

ACKNOWLEDGMENT

The authors acknowledge Beijing Jiaotong University forthe available Φ-OTDR dataset [16].

REFERENCES

1. Y. Fang et al., “Phase-sensitive optical time-domain reflectometry: A review,” IEEE Sens. J., vol. 19, no. 18, pp. 6548–6558, 2019.
2. J. Kang et al., “Automatic monitoring of rock-slope failures using distributed acoustic sensing and semi-supervised learning,” Geophys. Res. Lett., vol. 51, p. e2024GL110672, 2024.
3. A. Minardo et al., “Calibration of phase-sensitive optical time-domain reflectometry for distributed vibration sensing,” J. Lightw. Technol., vol. 40, no. 12, pp. 3723–3730, 2022.
4. R. Rayhana, G. Z. Xiao, and Z. Liu, “Fiber optic distributed sensing for pipeline integrity monitoring: Recent advances,” Opt. Fiber Technol., vol. 82, p. 103–115, 2025.
5. L. Cavuto et al., “Measurement accuracy and uncertainty in phase-sensitive OTDR systems,” Meas. Sci. Technol., vol. 31, no. 8, p. 085402, 2020.
6. Y. Li et al., “Evaluation of signal-to-noise ratio in Φ-OTDR distributed fiber sensing systems,” IEEE Trans. Instrum. Meas., vol. 70, pp. 1–12, 2021.
7. A. Tosi et al., “Thermal effects on laser frequency stability in coherent optical sensing,” Opt. Express, vol. 31, no. 14, pp. 22876–22889, 2023.
8. M. Garcia et al., “Calibration methods for long-range phase-sensitive optical time-domain reflectometry,” Appl. Opt., vol. 63, no. 5, pp. 1123–1134, 2024.
9. Z. Li et al., “Attention-based deep learning for noise suppression in distributed fiber sensing,” IEEE Photon. Technol. Lett., vol. 33, no. 18, pp. 985–988, 2021.
10. H. Tian et al., “Attention mechanisms for spatiotemporal feature extraction in Φ-OTDR,” Opt. Lett., vol. 46, no. 22, pp. 5563–5566, 2021.
11. Y. Shan et al., “Multidimensional convolutional neural networks for distributed acoustic sensing,” Neural Comput. Appl., vol. 36, pp. 1245–1258, 2024.
12. H. Tian et al., “Temporal modeling limitations of pure CNN architectures in Φ-OTDR signal processing,” IEEE Sens. J., vol. 22, no. 15, pp. 15234–15242, 2022.
13. Z. Liu et al., “KAN: Kolmogorov-Arnold Networks,” arXiv preprint arXiv:2404.19756, 2024.
14. I. Livieris et al., “A survey on Kolmogorov-Arnold Networks: Theory, variants, and applications,” Neurocomputing, vol. 592, p. 127874, 2024.
15. A. Vacarubio et al., “Optimization challenges of KANs on high-dimensional noisy data,” IEEE Access, vol. 12, pp. 88456–88467, 2024.
16. X. Cao, Y. Su, Z. Jin, and K. Yu, “An open dataset for phase-sensitive OTDR with machine learning applications,” Sci. Data, vol. 10, p. 789, 2023.

17. W. Cheng, Q. Zhang, S. Wen, B. Zhu, Q. Hu, and Z. Zhang, "Φ-GLMAE: Phase-sensitive global-local masked autoencoder for distributed acoustic sensing," IEEE Trans. Neural Netw. Learn. Syst., early access, 2025.
18. T. Das et al., "Progress in distributed fiber sensing: A review," J. Lightw. Technol., vol. 42, no. 5, pp. 1234–1256, 2024.
19. M. Tim et al., "Fiber optic sensor calibration for distributed measurement systems," IEEE Trans. Instrum. Meas., vol. 71, pp. 1–15, 2022.
20. T. Das et al., "A review of noise sources in distributed acoustic sensing," IEEE Sens. J., vol. 25, no. 3, pp. 3456–3468, 2025.
21. K. Metrologia et al., "Fiber optic gyroscope calibration and uncertainty evaluation," Metrologia, vol. 58, no. 4, p. 045012, 2021.
22. L. Measurement et al., "Measurement uncertainty in fiber optic sensing systems," Meas. Sci. Technol., vol. 31, no. 12, p. 124002, 2020.
23. H. Label et al., "Label-free anomaly detection in distributed fiber sensing," Opt. Express, vol. 31, no. 8, pp. 12345–12358, 2023.
24. Z. Hybrid et al., "Hybrid machine learning and deep learning for Φ-OTDR event recognition," IEEE Sens. J., vol. 21, no. 18, pp. 20456–20467, 2021.
25. Y. Dual et al., "Dual-stage recognition for distributed fiber sensing systems," Opt. Fiber Technol., vol. 68, p. 102834, 2022.
26. T. Das et al., "Review of phase-sensitive optical time-domain reflectometry," J. Opt., vol. 21, no. 9, p. 093001, 2019.
27. R. Semi et al., "Semi-supervised learning for distributed acoustic sensing," Neurocomputing, vol. 452, pp. 234–245, 2021.
28. Z. CNN et al., "CNN with attention for noise suppression in OTDR," IEEE Photon. Technol. Lett., vol. 33, no. 15, pp. 876–879, 2021.
29. H. TCN et al., "Temporal convolutional networks with attention for fiber sensing," Opt. Lett., vol. 46, no. 18, pp. 4567–4570, 2021.
30. Y. Multidim et al., "Multidimensional attention mechanisms for distributed sensing," IEEE Trans. Neural Netw. Learn. Syst., vol. 35, no. 6, pp. 7890–7902, 2024.
31. X. ResNet et al., "ResNet improvements for signal classification," IEEE Access, vol. 10, pp. 45678–45689, 2022.
32. W. ResNet et al., "Channel attention in ResNet for measurement applications," IEEE Trans. Instrum. Meas., vol. 74, pp. 1–12, 2025.
33. T. Das et al., "Calibration methods for distributed acoustic sensing," IEEE Trans. Instrum. Meas., vol. 73, pp. 1–15, 2024.
34. M. Measurement et al., "Calibration techniques for measurement consistency," Metrologia, vol. 61, no. 2, p. 025006, 2024.
35. P. MST et al., "Phosphor thermometry calibration standards," Meas. Sci. Technol., vol. 32, no. 7, p. 075103, 2021.
36. S. CNN et al., "Hybrid CNN-ResNet architectures for sensor data," IEEE Sens. J., vol. 25, no. 4, pp. 5678–5690, 2025.
37. A. KANet et al., "KANet: Kolmogorov-Arnold network for indoor inertial navigation," IEEE Trans. Instrum. Meas., vol. 74, pp. 1–14, 2025.
38. B. TCN-KAN et al., "TCN-KAN-FBM: Remaining useful life prediction with Kolmogorov-Arnold networks," IEEE Trans. Instrum. Meas., vol. 74, pp. 1–16, 2025.
39. L. Chen et al., "AttCWKAN: Attention-based continuous wavelet Kolmogorov-Arnold network for wind turbine fault diagnosis," IEEE Trans. Ind. Electron., vol. 72, no. 5, pp. 4567–4578, 2025.
40. C. KAN et al., "KANs for time series analysis: Stability and convergence," IEEE Access, vol. 12, pp. 98765–98778, 2024.
41. C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," J. Big Data, vol. 6, no. 1, p. 60, 2019.
42. T. Das et al., "Data augmentation strategies for fiber optic sensing systems," IEEE Sens. J., vol. 20, no. 15, pp. 8234–8245, 2020.

43. S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in Proc. ICML, 2015, pp. 448–456.
44. H. Time et al., “Time series decomposition using diffusion models for sensor data enhancement,” IEEE Trans. Neural Netw. Learn. Syst., vol. 36, no. 2, pp. 456–468, 2025.
45. Y. Fast et al., “FastGAN-based augmentation for time series classification,” Neurocomputing, vol. 610, p. 128456, 2025.
46. Z. Event et al., “Event-aware data augmentation for distributed acoustic sensing,” Opt. Express, vol. 32, no. 10, pp. 15678–15689, 2024.
47. K. He et al., “Deep residual learning for image recognition,” in Proc. CVPR, 2016, pp. 770–778.
48. S. Elfwing et al., “Sigmoid-weighted linear units for neural network function approximation in reinforcement learning,” Neural Netw., vol. 107, pp. 3–11, 2018.
49. D. Hendrycks et al., “Gaussian error linear units (GELUs),” arXiv preprint arXiv:1606.08415, 2016.
50. I. Loshchilov et al., “Decoupled weight decay regularization,” in Proc. ICLR, 2019.
51. J. Wang et al., “Accurate position measurement in phase-sensitive OTDR using advanced signal processing,” IEEE Trans. Instrum. Meas., vol. 73, pp. 1–12, 2024.
52. Y. Zhang et al., “Harmonic analysis for coherent noise suppression in distributed fiber sensing,” Opt. Express, vol. 32, no. 5, pp. 6789–6801, 2024.
53. H. Xiao et al., “Phase noise compensation techniques for Φ-OTDR systems,” J. Lightw. Technol., vol. 42, no. 8, pp. 2456–2468, 2024.
54. H. Li, C. Fan, T. Liu, and Z. Wang, “Deep learning-based event recognition in Φ-OTDR sensing systems,” Applied Optics, vol. 61, no. 11, pp. 2975–2997, 2022.
55. Y. Li, X. Zhang, and H. Wang, “Signal enhancement through data fusion of premium sensing channels in distributed fiber sensing,” Meas. Sci. Technol., vol. 36, no. 3, p. 035012, 2025.
56. S. Xie, J. Liu, and M. Chen, “Pulse intensity coding with mismatched filtering for phase-sensitive OTDR systems,” Opt. Express, vol. 32, no. 15, pp. 26789–26801, 2024.
57. X. Lan and X. Li, “STFT-AECNN: An attention-enhanced CNN for efficient Φ-OTDR event recognition in IoT-enabled distributed acoustic sensing,” arXiv preprint arXiv:2509.19281, 2025.