

ARTICLE

Real-Time Fault Diagnosis in UHV Power Systems Using Bayesian-Optimized 1D CNN on Raw Time-Series Signals

Hani Albalawi^{1,2}, Rab Nawaz³, Abdul Wadood^{1,2,*}, Anwar Ul Haq³, Shahbaz Khan^{1,2},
Bakht Muhammad Khan¹ and Aadel Mohammed Alatwi^{1,2}

¹Zero Emission Technologies Innovation Center, University of Tabuk, Tabuk, Saudi Arabia

²Electrical Engineering Department, Faculty of Engineering, University of Tabuk, Tabuk, Saudi Arabia

³Department of Electrical Engineering, Mirpur University of Science and Technology, Mirpur, Azad Jammu and Kashmir, Pakistan

*Corresponding Author: Abdul Wadood. Email: wadood@ut.edu.sa

Received: 28 April 2026; Accepted: 24 August 2026; Published: 28 September 2026

ABSTRACT: Ensuring the reliability of Ultra-High Voltage (UHV) power systems remains a critical challenge, as series compensation improves stability while introducing complex fault dynamics. Although machine learning and deep learning methods have advanced fault diagnosis, existing approaches often depend on computationally intensive preprocessing and struggle with data scarcity, limiting real-time applicability. This study proposes a streamlined One-Dimensional Convolutional Neural Network (1D CNN) optimized via Bayesian learning for efficient and robust fault classification in a 735 kV, 32-bus UHV system. The model operates directly on raw time-series signals, eliminating the need for domain-specific transformations while preserving the natural characteristics of the data. Quantitatively, the proposed framework achieves 99.91% test accuracy using combined voltage-current signals, with a narrow standard deviation of $\pm 0.06\%$ across multiple evaluations, and correctly classifies 12,998 out of 13,010 test instances with only twelve misclassifications across eleven fault categories. For resource-constrained applications, voltage-only signals deliver 99.72% accuracy while reducing training time by 38.6% (680 s) and memory usage by 61.8% (1272 MB), whereas current-only signals achieve 99.90% accuracy with the fastest inference speed of 2995 samples per second. The model demonstrates strong dependability under noisy conditions, maintaining 97.03% test accuracy even at 10 dB SNR, and exhibits robustness against downsampling with a Sampling Index of 0.0235. The Bayesian optimization process identifies optimal hyperparameters including a learning rate of 0.002611, dropout rate of 0.182, and batch size of 19. From an economic perspective, the elimination of expensive preprocessing, noise cancellation, and high-frequency sampling (typically 0.2–10 MHz) reduces computational resource requirements and operational costs, enabling practical deployment in real-time monitoring systems. These findings demonstrate the practical potential of the proposed approach for cost-effective, scalable fault diagnosis in large-scale power grids.

KEYWORDS: Attention mechanism; Bayesian optimization; computational efficiency; performance generalization; series-compensation; time series analysis

1 Introduction

High- Voltage UHV power systems require series compensation to ensure system stability and reliability. However, the operating conditions become more complex and difficult to manage due to the interactions caused by the series compensation [1,2]. Among the abnormal conditions in power system operation, the impact of short-circuit faults (SCF) is the most severe. If the impact of SCF on the power system is not properly mitigated, it could lead to catastrophic failures, incurring considerable economic costs for repair and

replacement [3]. Machine learning and deep learning methods have been proposed to address the limitations in the development of robust fault diagnosis methods [4–7].

Table 1 lists the abbreviations and acronyms used throughout this paper for quick reference.

Table 1: List of abbreviations.

Abbreviation	Description	Abbreviation	Description
1D CNN	One-Dimensional CNN	MODWPT	Maximal Overlap DWT
BOA	Bayesian Optimization Accuracy	MSE	Mean Square Error
CSA	Convolutional Sparse Autoencoder	NII	Noise Immunity Index
CNSF	Capsule Network with Sparse Filtering	ORNL	Oak Ridge National Laboratory
CWT	Continuous Wavelet Transform	PMU	Phasor Measurement Unit
DG	Distributed Generation	PSO	Particle Swarm Optimization
DWT	Discrete Wavelet Transform	PQD	Power Quality Disturbance
EMD	Empirical Mode Decomposition	ReLU	Rectified Linear Unit
FC	Fault Classification	RF	Random Forest
FDC	Fault Detection and Classification	RFE	Recursive Feature Elimination
FDCL	Fault Detection, Classification and Localization	RNN	Recurrent Neural Network
FL	Fault Localization	SAT-CNN	Self-Attentive CNN
FREEDM	Future Renewable Electric Energy Delivery and Management	SCADA	Supervisory Control and Data Acquisition
GNN	Graph Neural Network	SCF	Short-Circuit Fault
GTO	Gorilla Troops Optimizer	SDA	Stacked Denoising Autoencoder
GWO	Grey Wolf Optimizer	SI	Sampling Index
HIF	High-Impedance Fault	SMOTE	Synthetic Minority Over-sampling
HSA	Harmony Search Algorithm	SNR	Signal-to-Noise Ratio
ID	Intrusion Detection	STATCOM	Static Synchronous Compensator
LIF	Low-Impedance Fault	STFT	Short-Time Fourier Transform
LightGBM	Light Gradient Boosting Machine	TCSC	Thyristor-Controlled Series Capacitor
LOLO-CV	Leave-One-Loading-Out CV	TW	Traveling Wave
LSTM	Long Short-Term Memory	UHV	Ultra-High Voltage
MAE	Mean Absolute Error	VSC	Voltage Source Converter
MAPE	Mean Absolute Percentage Error	WELM	Weighted Extreme Learning Machine
MMC	Modular Multi-Level Converter	WGN	White Gaussian Noise

One of the main challenges in using the concept of data-driven artificial intelligence (AI) in fault diagnosis lies in the scarcity of the data itself, owing to the constraints on the availability of fault data in the real world due to privacy and security concerns. Hence, standardized simulation models, which have been well-accepted in the power system field, offer a controlled environment to generate a wide range of fault conditions, thereby providing comprehensive fault data considering the diversity in fault types and system complexities for the purpose of validation [8].

The proposed study is developed and validated using a MATLAB/Simulink-based high-fidelity 735 kV, 32-bus UHV transmission grid system [9]. In power system research, real-time measurements are often scarce and not publicly available owing to security and privacy concerns related to critical infrastructure. Access to hardware-in-the-loop platforms and dedicated real-time simulators is limited in several research environments. Therefore, simulation-based frameworks are widely accepted and provide a controlled and reproducible environment for systematic fault diagnosis analysis. Although this study does not employ hardware-in-the-loop or physical real-time simulators, the MATLAB/Simulink-based UHV system captures practical power system dynamics and supports development and evaluation of advanced deep learning frameworks with potential real-time applicability. Here, “real-time” refers to the proposed 1D-CNN model’s ability to perform fast inference on raw time-series data in simulation without computationally expensive signal processing. The framework remains consistent with established practices and supports reproducibility and future real-time deployment. Although various methodologies [10–12] have been shown to have high accuracy and fast learning capabilities, end-to-end processing is essential. This approach relies on the One-Dimensional Convolutional Neural Network (1D CNN) and evaluates the accuracy, computation time, and memory consumption, which are very critical factors in the development of current technologies.

Most of the previously used techniques [13,14] prove to be accurate enough in fault diagnosis, but they highly rely on the preprocessing stages that involve image transformation techniques, making them impossible to use in real-time monitoring because of the complexity of calculations. On the other hand, the suggested technique provides an extensive analysis of computation time at each step of the processing procedure [15–17], which is usually omitted in previous works [18].

Classical and advanced optimization classifiers [12,19–22] rely on handcrafted feature engineering for effective learning, which requires domain expertise and may discard discriminative temporal and spectral information embedded in raw fault signals. As power system complexity increases, designing features that generalize across diverse fault scenarios can become increasingly challenging, limiting model scalability and robustness. Moreover, hybrid fault ride-through strategies can improve the overall resilience and operational feasibility of next-generation power systems such as HVDC transmission systems [23].

Although the methodologies [24–26] show efficient model learning using numerical or transformed data from Simulink standard models, further optimization in the feature extraction phase is necessary to ensure the efficiency of the end-to-end processing. Moreover, although the proposed methodologies [1,15,27] show high noise immunity, they use noise cancellation techniques, which increase the computational cost. However, the existing methodologies face difficulties in processing high noise levels of 10–20 dB [28–30]. Improving noise immunity may require sophisticated noise-cancellation or denoising techniques, which can increase computational complexity and hinder real-time deployment. The proposed work uses the nominal sampling frequency and investigates the impact of the sampling frequency on fault signal processing. Moreover, the proposed work eliminates the need for high sampling frequencies [8,31], which usually range from 0.2 to 10 MHz. Prior research [19,27,32,33] indicates combined voltage-current (VI) signals outperform individual voltage or current signals, though VI signals remain susceptible to downsampling effects, underscoring the need for optimal sampling frequency selection.

For optimizing the model learning, Bayesian optimization (BO) is employed, which is more efficient than traditional grid and random search methods [34]. Although the efficiency of some techniques has been shown to improve the performance of fault diagnosis, such as the use of harmony search and particle swarm optimization algorithms for parameter tuning and classifier co-training, it has been argued that these approaches are computationally costly and there exists a trade-off [26,35,36]. In the proposed research study, the scalability and adaptability of the fault diagnosis system will be assessed by taking into consideration a large-scale UHV power system with compensation systems. This is because scalability and adaptability are important considerations that enable the fault diagnosis system to handle increasing data patterns and fault complexities.

The proposed methodology overcomes the challenges in fault diagnosis in large-scale series-compensated UHV power systems by utilizing the proposed 1D CNN model with Bayesian optimization. This facilitates efficient processing without the need for intensive pre-processing. The computational resources required for the proposed approach are analyzed to provide a balance between efficiency and accuracy. The simulation models employed in the suggested method enable thorough validation of the approach. This method enhances adaptive fault diagnosis systems for UHV power systems.

1.1 Contribution

The study makes the following key contributions for large-scale UHV power systems:

- End-to-end learning with a 1D CNN enables the autonomous extraction of optimal features from raw fault signals in time-series form, eliminating the need for manual feature engineering.
- The convolutional model minimizes memory requirements and achieves consistent execution speed irrespective of system complexity, thus improving the efficiency of computation.
- Inclusion of Bayesian optimization enables efficient model learning compared to conventional parameter tuning techniques, thus improving the efficiency of the model.
- In-depth analysis of the efficiency of computation time, memory usage, noise immunity, and sampling rate effects enables a scalable approach for real-time fault diagnosis in large series-compensated UHV systems.

1.2 Literature Review

In [1], a Self-Attentive Weight-Sharing Capsule Network (WSCN) has been proposed to ensure accurate fault diagnosis in series-compensated transmission lines using robust classification with limited data and highlighting the features using weight sharing. In [2], a lightweight CNN fine-tuned with energy valley optimizer was designed using Continuous Wavelet Transform (CWT) scalograms, achieving superior fault diagnosis performance compared to benchmark architectures on balanced, unbalanced, and noisy datasets. This research in [3] introduced an RF-LSTM-tuned KNN ensemble that achieved 99.96% accuracy for fault detection and classification by analyzing voltage and current patterns across transmission line phases.

In [4], empirical mode decomposition extracts features from voltage and current waveforms to train an Extreme Gradient Boost classifier, achieving 97.59% accuracy for incipient fault identification in underground cables. In [5], an Autogluon-based machine learning method for fault diagnosis in transmission lines showed how important feature engineering and the Minority Over-sampling Technique (SMOTE) data balance were to improving diagnostic accuracy. In [6], a graph autoencoder-based detector with bi-level optimization was developed for false data injection attacks on voltage regulation, achieving a 98.11% detection rate on the 486-bus Iberian power system.

In [7], three deep recurrent neural network (RNN) models using Long Short-Term Memory (LSTM) and Phasor Measurement Unit (PMU) data were introduced for fault region identification, type classification, and location prediction, achieving superior accuracy in a two-area four-machine system. In [8], a traveling wave (TW)-based fault location method was proposed combining time- and frequency-domain features, achieving low estimation error across the IEEE 34-node system feeder length.

In [10], a Wide Neural Network (WNN) was presented using PMU voltage, current magnitude, and phase angles for fault location on the Western Systems Coordinating Council (WSCC) 9-bus system, achieving low prediction errors.

This article [11] proposed a fault classification algorithm using Maximal Overlap Discrete Wavelet Packet Transform (MODWPT) and a bagged tree ensemble for Thyristor-Controlled Series Capacitor (TCSC)-compensated transmission lines, achieving 100% accuracy on the WSCC 9-bus system. In [12], a Particle Swarm Optimization (PSO)-weighted ensemble method was presented for real-time fault detection using raw data, validated on IEEE 14-bus and 39-bus systems without pre-computational techniques. The study [13] proposed a CNN with Gorilla Troops Optimizer (GTO) adaptive protection scheme for the Future Renewable Electric Energy Delivery and Management microgrid, achieving 99.37% fault detection, 99% classification, and 98.2% location accuracy.

In [14], a data reduction method using Short-Time Fourier Transform (STFT) and Fast-DTW was proposed to reduce redundant data by 40.2% before CNN training. The approach achieved 99.37% fault classification accuracy in simulations. In [15], SAT-CNN (self-attention CNN) along with time-series images was introduced to classify faults in transmission lines. This model gave high accuracy in the presence of both voltage and current data at different sample rates.

A machine learning approach for transmission line fault detection and location has been proposed in [16]. XGB obtained a classification accuracy of 99.82%, whereas the MAPE of CNN-LSTM is less than 1% and the mean absolute error (MAE) is less than 0.16 km. In [17], a deep learning method using a CNN was proposed for power quality disturbance (PQD) diagnosis. It classified fault types and locations with over 99% accuracy using 50 Hz simulated data.

In [18], three deep learning models (CNN, LSTM, and CNN-LSTM) were proposed for smart grid fault diagnosis on IEEE 6-bus and 9-bus systems. They outperformed existing methods in detection, classification, and localization accuracy. In [19], four ensemble classifiers were evaluated for fault diagnosis on series-compensated power transmission lines (SC-PTL). RF achieved 99.94% accuracy, while XGB showed the fastest computation at 0.2790 s. In [20], an ML-based intrusion detection framework was proposed for Supervisory Control and Data Acquisition (SCADA) power systems using the Oak Ridge National Laboratory (ORNL) dataset. RF achieved a 94.09% F1 score on unseen data after augmentation and balancing.

In [21], eight classifiers were evaluated for fault classification using a standardized Simulink dataset. Logistic Regression (LR), RF, and Support Vector Machines (SVM) outperformed others with high accuracy and reduced computational time. In [22], A hybrid framework combining 21 domain-knowledge features with machine learning was proposed for transmission line fault diagnosis. XGB achieved 94.25% accuracy across 11 fault categories. In [24], an SA-MobileNetV3 model with CWT-based 2D image conversion was proposed for transmission line fault classification. It achieved 99.90% accuracy across 11 fault types.

In [25], Power System Machine Learning (PSML), an open-access multi-scale time-series dataset for joint transmission and distribution grids, was presented to enable ML-based reliable grid operation and decarbonization research. In [27], A SAT-CNN with DWT denoising and time-series imaging was proposed for transmission line fault classification. It achieved high accuracy and noise immunity. In [28], a transfer

learning method using a Stacked Denoising Autoencoder (SDA) was proposed for Voltage Source Converter High-Voltage Direct Current (VSC-HVDC) fault location. It performed effectively with small datasets under various fault conditions.

In [29], fuzzy thresholding, machine learning with wild horse optimization, and adaptive neural fuzzy inference were proposed for smart grid fault diagnosis. RF achieved 100% accuracy with 3.54×10^{-6} Mean Squared Error (MSE). In [30], a microgrid protection scheme using MODWT feature extraction and XGB classification was proposed. It achieved effective fault detection and classification under varied configurations and conditions. In [31], Independent Component Analysis (ICA), TW theory, and SVM were combined for fault location and classification in HV transmission lines. It achieved <1% error and 100% accuracy under noisy conditions.

In [32], a capsule network with sparse filtering (CNSF) was proposed for unsupervised fault detection and classification in transmission lines. It performed effectively under topology changes, noise, and high-impedance faults (HIF). In [33], a Convolutional Sparse Autoencoder (CSA) was proposed for transmission line fault detection and classification. The system automatically learned its features and showed good robustness and generalizability. According to [34], an LSTM model along with DWT-based feature extraction was used to find HVDC faults. This method was capable of achieving an accuracy of 99.04% with a fault resistance of 480 ohms and using a three-level relay system

The hybrid method using ANN along with RF with the help of the Optuna optimization technique was proposed in [36] for fault diagnosis in electrical power networks. Using Synthetic Minority Over-sampling Technique (SMOTE)-balanced data, the model was able to obtain an accuracy of 99.8%, outperforming other traditional and advanced classifiers. However, challenges exist regarding computational efficiency and practical validation. In [37], a unified multi-task learning framework with an attention mechanism and Optuna optimization achieved high-accuracy fault identification, classification, and localization on the IEEE 39-bus system, outperforming traditional models. In [38], an intelligent relaying scheme was proposed using Variational Mode Decomposition (VMD)-extracted current signal features and a CNN classifier for fault diagnosis in Doubly-Fed Induction Generator Unified Power Flow Controller (DFIG-UPFC) transmission lines, achieving fast fault detection (<10 ms) and 99.86% classification accuracy.

A robust CNN-BFO architecture was proposed in [39] for PV fault identification based on the integration of five CNNs (GoogleNet, SqueezeNet, ResNet-50, VGGNet-16, and AlexNet) and Bitterling Fish Optimization (BFO) for selection. VGGNet-16+BFO achieved 98.75% accuracy, 98.72% sensitivity, 98.78% specificity, 98.76% precision, and 98.74% F1 score, outperforming other meta-heuristic optimizers. The study in [40], performed fault analysis of UHV power systems via simulation-based fault diagnosis. The methods of feature engineering and ensemble learning were used to get accurate results for fault detection and localization.

The issue of fault detection in series-compensated ultra-high voltage transmission lines continues to be difficult in spite of notable progress made through the use of artificial intelligence. Earlier studies have demonstrated classification accuracies higher than 99% on a wide range of power systems of different levels of complexity. However, there are some constraints associated with these methods, which restrict their implementation in real-time applications to large-scale series-compensated power systems. These methods utilize many complex pre-processing algorithms such as HSA, DWT, CWT, EMD, MODWPT, and image-based signal processing methods. For instance, the hybrid machine and deep learning frameworks reported in [3,16,40] required processing times of 10,920, 2712.80, and 840 s, respectively, to achieve optimal performance. These substantial computational requirements, together with significant memory demands, may limit their suitability for real-time protection applications.

Moreover, the test and validation of existing methodologies in small or medium-sized systems like IEEE 4-bus, 6-bus, 9-bus, and 14-bus systems, without using series compensation and operating at UHV levels, does not show the scalability of these methodologies under real-world system scenarios. The reliance on engineered features and transformation from signal to images can lead to a loss of valuable time data present within the transient fault signals, as well as increased computational overhead, rendering it impractical for implementation. Signal ablation studies have not been sufficiently carried out, resulting in an inability to determine the efficiency of diagnosis if partial data from sensors is unavailable. Scalability in terms of sensitivity to various sampling rates from 1 to 100 MHz and signal-to-noise ratio of 20 dB shows performance deterioration to around 97% precision. Many current methodologies do not adequately report inference times, training times, memory requirements, and their ability to generalize fault classifications for multiple fault resistances and series-compensated networks.

Therefore, these challenges highlight the critical necessity of an efficient and effective end-to-end fault diagnostic model able to process unprocessed multivariate discrete-time data directly without any computationally expensive feature extraction step. This model has to achieve accurate diagnostics while significantly lowering the preprocessing costs, making sure to attain efficient and effective training and inference with the possibility of implementing real-time operation, and at the same time exhibiting good performance even under signal masking and noise injection, as well as varying sample rates, in large series-compensated UHV transmission systems. In order to tackle the aforementioned problems, this study introduces a Bayesian-optimized One-Dimensional Convolutional Neural Network fault diagnostic framework, which learns discriminative features directly from the raw signal measurements with emphasis on accurate fault diagnostics, efficient computation, robustness, and potential real-time deployment.

1.3 Work Organization

This paper is organized as follows: The next [Section 2](#) discusses the proposed framework for the Bayesian-optimized one-dimensional convolutional neural network. This includes dataset pre-processing or signal de-interleaving, a one-dimensional convolutional neural network for feature extraction, classification and optimization, performance metrics, and lastly the dataset used for the experiment. The experimental findings and their corresponding discussion are presented in [Section 3](#). This includes accuracy and efficiency analysis, Bayesian optimization and learning, performance generalization and overfitting characteristics, noise analysis, sampling analysis, the perfect recipe or the optimized hyperparameters identified by the Bayesian optimization algorithm, comparison with existing work, and lastly the future scope. The last [Section 4](#) discusses the conclusion.

2 Proposed Bayesian-Optimized 1D CNN Framework

The proposed methodology for multi-class fault classification involves three main stages: (1) Dataset structuring via signal de-interleaving, (2) Model development using a Bayesian-optimized 1D Convolutional Neural Network (CNN), and (3) Performance evaluation.

[Fig. 1](#) shows a fault classification pipeline using a 1D CNN with leakage prevention. Raw signals are preprocessed to a fixed length before train/test splitting. Training involves normalization on the training dataset; Bayesian optimizations are done only using the training dataset, with the hyperparameters optimized using standard 10-fold cross-validation and hyperparameter evaluation. Convolution, pooling, depthwise blocks, global attention, and softmax functions are used in the model for classification purposes. Test datasets are kept separate with normalized values and one-off evaluations. Results include accuracy, confusion matrix, and classification report, maintaining strict train-test separation while optimizing performance.

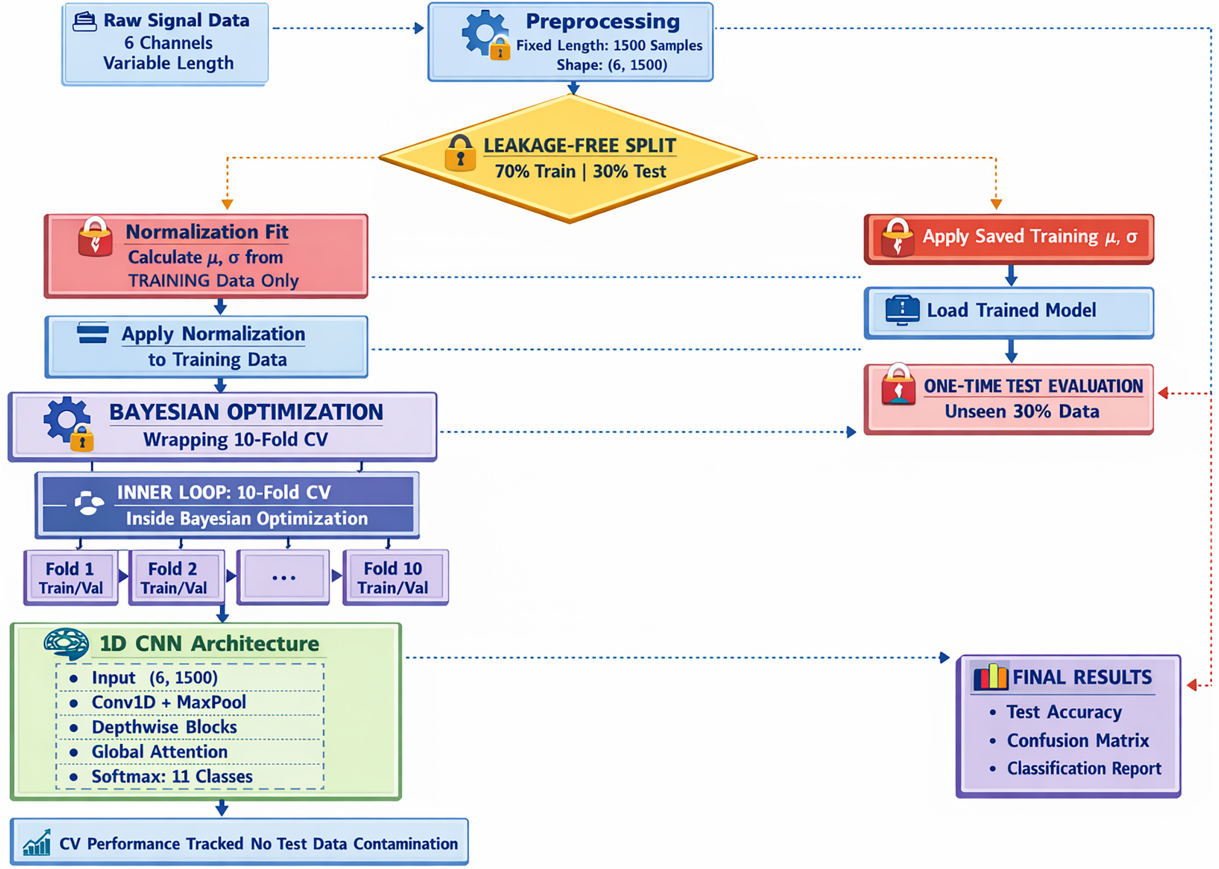


Figure 1: Leakage-aware 1D CNN fault classification pipeline with Bayesian optimization and 10-fold cross-validation in a series-compensated power transmission grid system.

2.1 Dataset Preprocessing: Signal De-Interleaving

The raw dataset is a composite matrix, $\mathbf{D}$, formed by vertically stacking individual instances. Each instance consists of interleaved fault signals (voltages and currents) concatenated with its corresponding class label. To prepare the data for the 1D CNN, the signals must be separated from the labels and restructured.

Let the raw data matrix be $\mathbf{D} \in \mathbb{R}^{M \times (L+1)}$, where M is the total number of instances, and L is the total length of the interleaved signal. The last column represents the class label $y \in \{0, 1, \dots, 10\}$ for 11 classes (10 fault types + 1 normal condition).

For a system with N recording points, each recording point contributes V variables (e.g., voltage and current, so $V = 2$). The total signal length per instance is $L = N \times V \times T$, where T is the number of time samples per variable.

De-interleaving is the process of reshaping each row of signals into a structured 2D format suitable for a 1D CNN. Specifically, a raw signal vector $\mathbf{x}_{raw} \in \mathbb{R}^L$ for a single instance is reshaped into a matrix $\mathbf{X} \in \mathbb{R}^{N \times (V \cdot T)}$. This matrix arranges the data so that each of the N rows includes the time-series information for all V variables from a single recording location. For a 1D CNN designed to process multi-variable time series, this matrix is often further transformed or treated as a multi-channel input of length N with $(V \cdot T)$ features, or alternatively, reshaped into a 1D tensor of length L . The labels are extracted to form the label vector $\mathbf{y}$.

2.2 1D Convolutional Neural Network for Feature Extraction

A 1D CNN is employed to automatically learn salient features from the de-interleaved fault signals. The core operation is the 1D convolution. Given an input signal f and a kernel g , the output feature map at layer l is given by:

$$\text{Feature}_l(i) = (f * g)(i) = \sum_{m=0}^{K-1} f(i+m) \cdot g(m) \quad (1)$$

where K is the kernel size. This operation is followed by a non-linear activation function. The Rectified Linear Unit (ReLU), defined as $\text{ReLU}(x) = \max(0, x)$, is used to introduce non-linearity.

To reduce dimensionality and extract dominant features, a max-pooling layer is applied. For a pooling size P and stride S , the output is:

$$\text{Pooled}(j) = \max_{n=0}^{P-1} \text{Feature}(j \cdot S + n) \quad (2)$$

After several convolutional and pooling layers, the extracted high-level features are flattened into a 1D vector $\mathbf{v}$ and passed through a series of fully connected (dense) layers.

2.3 Classification and Optimization

The final layer of the network is a dense layer with 11 neurons, employing the softmax activation function to produce a probability distribution over the 11 classes. For an input feature vector $\mathbf{v}$ and weights $\mathbf{W}$ and biases $\mathbf{b}$ of the final layer, the probability for class i is:

$$P(y = i|\mathbf{v}) = \text{softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=0}^{10} e^{z_j}}, \quad \text{where } \mathbf{z} = \mathbf{W}\mathbf{v} + \mathbf{b} \quad (3)$$

The model is trained by minimizing the categorical cross-entropy loss. For a single training sample with true class label y (one-hot encoded as $\mathbf{y}_{true}$), the loss $\mathcal{L}$ is:

$$\mathcal{L} = - \sum_{i=0}^{10} y_{true}^{(i)} \log(P(y = i|\mathbf{v})) \quad (4)$$

2.4 Hyperparameter Optimization Using Bayesian Optimization

To automate and optimize the network's architecture and training hyperparameters (e.g., number of filters, kernel size, learning rate, number of layers), Bayesian Optimization is employed. The aim is to identify the hyperparameter configuration λ^* that optimizes a specified objective function, which is usually the cross-validation accuracy, $f(\lambda)$.

$$\lambda^* = \arg \max_{\lambda \in \Lambda} f(\lambda) \quad (5)$$

Bayesian optimization uses a probabilistic model of the objective function, typically a Tree-Structured Parzen Estimator or a Gaussian Process. It uses an acquisition function for optimization, which efficiently searches for the optimal set of hyperparameters by intelligently selecting the next set of hyperparameters to evaluate, thus being more efficient than grid or random search.

2.5 Performance Metrics

The optimized model's performance is assessed through various metrics:

- **Cross-Validation Accuracy:** The average accuracy obtained using the 10-fold cross-validation technique over the training data, which is used as the objective function $f(\lambda)$ in the optimization process.
- **Test Accuracy:** The accuracy of the trained model over the test data.
- **Training Time:** The time taken to train the model with the optimized hyperparameters.
- **Bayesian Optimization Time:** The time taken to finish the hyperparameter optimization process.
- **Memory Usage:** The memory used during the training process.
- **Classification Report and Confusion Matrix:** Calculated over the test data to obtain the precision, recall, and F1-score for each of the 11 classes (10 fault types + 1 normal condition), along with the visualization of the confusion matrix.

Analysis of the sampling frequency is important in fault classification since the fidelity of the signal is dependent on the sampling rate; failure to sample sufficiently can cause high-frequency fault information to be lost. The test setup produces multivariate fault signals having a sampling frequency of 20 kHz. In order to study the influence of the sampling frequency on fault detection, the signal is down-sampled from 1 to 15 kHz. This down-sampling is done under different operating conditions and fault transitions in order to systematically analyze the influence of sampling frequency on fault detection.

Noise tolerance is important in simulated fault detection since noise can affect the signals of power systems in the process of measuring, communicating, and through sensors. Simulating the model under various noise conditions tests how well the model performs. White Gaussian noise is used to simulate signals with SNRs varying between 10 and 50 dB.

The discrete-time fault signal $F[t]$ corrupted with WGN is given by (6)

$$F_{\text{noisy}}[t] = F[t] + n[t] \quad (6)$$

In this context, $n[t] \sim \mathcal{N}(0, \sigma^2)$ represents a Gaussian distribution with a mean of zero and a variance σ^2 , which is selected based on the desired signal-to-noise ratio (SNR). The SNR, expressed in decibels, is calculated as follows:

$$\text{SNR}_{\text{dB}} = 10 \cdot \log_{10} \left(\frac{\sum_t F[t]^2}{\sum_t n[t]^2} \right) \quad (7)$$

The process ensures that the noise introduced corresponds to the given SNR, thus enabling a systematic analysis of fault detection under noisy conditions.

2.6 Datasets

The experimental setup, as depicted in Fig. 2, is implemented in Simulink, which simulates a modern power transmission grid at 735 kV, along with a Modular Multi-Level Converter (MMC)-based STATCOM. This consists of 22 sub-modules per arm, which are delta-connected via an 18/735 kV transformer. By adjusting the voltage level of the converter relative to the grid, one can achieve different modes: the inductive mode occurs when the voltage is lower, whereas the capacitive mode happens when the voltage is higher.

The simulated system represents a 60 Hz, 32-bus network with six generating stations, multiple transformers, shunt elements, and 17 distributed parameter transmission lines, of which 11 are series-compensated (15%–42%) to enhance the power transfer capability. A 50 MVA STATCOM is included for

Figure 1: Power system diagram of the IEEE 30 bus system. The diagram is divided into two zones: North Zone (left) and South Zone (right). The North Zone includes a 5500 MVA generator, a 2200 MVA generator, and a 30 MVA STATCOM. It features several transmission lines, including SC-PTLs (Series-Compensated Power Transmission Lines) and DP-PTLs (Distributed Parameter Power Transmission Lines). The South Zone includes a 5600 MVA generator, a 5000 MVA generator, and a 280 MW load. It also features SC-PTLs and DP-PTLs. The diagram shows the interconnection between the two zones via a 50 km DP-PTL line. The legend indicates that DP-PTLs are represented by blue lines and SC-PTLs by red lines.

For data generation, simulations are conducted for 0.5 s at a 20 kHz sampling frequency (corresponding to 5×10^{-5} s sampling step), which aligns with many recent studies [1,13,19,36] that employ simulation setups for fault diagnosis in power systems. Fault signals $F[t]$ are extracted within the interval $(t_2 - t_1)$ corresponding to fault initiation and clearance. Fault signals are also preprocessed to remove the linearity effect and emphasize non-linearity among the variables to effectively capture the intricate interaction due to the complex power system configuration. The model uses different model parameters created from unique random numbers so that no repeated data is used. Such an approach reflects different power system conditions, such as high impedance fault, equipment saturation, oscillation, and inductor-capacitor switching [1]. Unlike methodologies [11,27] relying on discrete or limited parameter values, the proposed approach improves data quality by accurately simulating complicated cases. Moreover, the proposed analysis covers all the transmission lines and considers crucial power system nodes that are frequently ignored in the literature, following standard power engineering practice [41]. The location of signal recording devices and the recording approach are crucial in obtaining realistic data from the current power system [13]. By recording fault signals from 15 locations, the methodology captures large-scale system behavior and addresses the increasing complexity of diverse fault types and sizes. This framework ensures a representation of system dynamics under fault conditions.

Furthermore, faults are applied on the entire network, including transmission lines and buses, enabling system-wide analysis. This strategy improves the overall dataset diversity and can be effective in assessing the scalability of the proposed methodology based on the complexity of faults in large-scale UHV power systems. The fault signals are collected from all the main system buses to obtain fault data. This data recording scheme is essential for efficient fault diagnosis within power systems since the placement of PMUs

in a power system is significant for efficient analysis, as discussed in [10]. Table 2 shows the parameters used for simulation purposes. The parameters are varied based on randomized non-repeating numbers for efficient simulation results. This strategy ensures unique fault simulation (no duplications) based on dynamic operating conditions within power systems.

Table 2: Test system parameters are varied to generate simulation datasets. Parameters are controlled to avoid value repetition and different fault types simulate real-time conditions.

Parameter	Range
Fault type ¹	1–11
Fault resistance	0.1–100 Ω
Ground resistance ²	0.1–200 Ω
Fault duration	0.1–0.4 s
Fault inception Angle ³	0–0.01667 s
Fault location	Entire lengths of transmission lines, including the system nodes

Note: ¹All ten types of faults and the normal operating condition. ²Low-impedance fault (LIF) and high-impedance faults (HIF) are included. ³Fault can occur at any instant of the complete 60 Hz AC waveform.

Algorithm 1 outlines the process of creating a labeled dataset of short-circuit fault signals with the help of a Simulink UHV power grid model by varying the fault parameters randomly over N iterations. The implementation at every iteration is sampling various non-repeating values for fault timing (τ_1, τ_2), resistance (r_f, r_g), location (ℓ), and type (y of 11 configurations), then setting up and running a 0.5 s simulation. The fault inception (inception angle) listed in Table 2 (0–0.01667 s) represents the point within the 60 Hz AC cycle where the fault is initiated. Since the fault start time τ_1 is randomly selected from 0.1 to 0.4 s, the fault can occur at any arbitrary instant within the simulation window. Consequently, the fault inception angle is implicitly random and covers the entire AC cycle (0–0.033 s). The dynamic fault duration is computed as $\tau_2 - \tau_1$, with a minimum duration of 0.07 s enforced to ensure sufficient transient data capture. The generated 6-channel signal (three voltages, three currents) is recorded within the fault interval $[\tau_1, \tau_2]$, transformed into a row vector, and the fault type label is added in front.

Algorithm 1: Generation of short-circuit fault signals in UHV power grid system

Require:

- 1: **Simulink model:** power-grid-statcom
 - 2: **Transmission line parameters, fault type set** $F = \{1, 2, \dots, \kappa\}$, **where** $\kappa = 11$
 - 3: **Parameter ranges:** fault resistance $R_f \in [0.1, 100] \Omega$, ground resistance $R_g \in [0.1, 200] \Omega$, fault location $d \in [d_{min}, d_{max}]$ km, d_{min}, d_{max} depends upon the length of the transmission line
 - 4: **Sampling frequency** $f_s = 20$ kHz, **simulation duration** $T_{sim} = 0.5$ s
- Ensure: A dataset D of labeled multi-variate discrete-time fault signals $\{X_i\}_{i=1}^N$**
- 5: Initialize empty dataset $D \leftarrow 0$
 - 6: Set maximum time steps $T_{max} \leftarrow 0$
 - 7: **for** $i = 1$ to N **do**
 - 8: **Generate non-repeating parameters**
 - 9: Generate random fault start time $\tau_1 \sim U(0.1, 0.4)$
 - 10: Generate fault inception $t_{inc} \leftarrow \text{mod}(\tau_1, 1/60)$
 - 11: Generate random fault end time $\tau_2 \sim U(0.1, 0.4)$
-

(Continued)

Algorithm 1 (continued)

```

12:   while  $\tau_2 - \tau_1 < 0.07$  do
13:     Resample  $\tau_1, \tau_2 \sim U(0.1, 0.4)$ 
14:   end while
15:   Compute fault duration  $d_f \leftarrow \tau_2 - \tau_1$  {Fault inception within AC cycle (00.01667 s for 60 Hz)}
16:   Generate random fault resistance  $r_f \sim U(0.1, 100)$ 
17:   Generate random ground resistance  $r_g \sim U(0.1, 200)$ 
18:   Generate random fault location  $\ell \sim U(d_{min}, d_{max})$  km
19:   Select fault type  $y \sim U\{1, \dots, \kappa\}$ 
20:   Eleven distinct fault configurations
21:   Configure Simulink model parameters:
22:     Set SwitchTimes  $\leftarrow [\tau_1, \tau_2]$ 
23:     Set FaultResistance  $\leftarrow r_f$ 
24:     Set GroundResistance  $\leftarrow r_g$ 
25:     Set transmission line lengths based on  $\ell$ 
26:     Configure breaker states according to  $y$ 
27:   Execute Simulink simulation  $\text{sim}(\text{'power-grid-statcom'}, T_{\text{sim}})$ 
28:   Acquire multivariate signal  $S \in R^{T \times Q}$ , where  $Q = 6$ 
29:     where columns  $\{v_a, v_b, v_c\}$  represent phase voltages
30:     and columns  $\{i_a, i_b, i_c\}$  represent phase currents
31:   Extract fault segment  $F_i \leftarrow S(t \in [\tau_1, \tau_2], :)$ 
32:   Let  $T_i$  be the number of time samples in  $F_i$ 
33:   Interleave multi-variate channels
34:   Construct row vector  $x_i \leftarrow \text{vec}(F_i^T)^T$   $\{x_i \in R^{1 \times QT_i}\}$ 
35:   Prepend fault label:  $\tilde{x}_i \leftarrow [y, x_i]$ 
36:   Update  $T_{\max} \leftarrow \max(T_{\max}, T_i)$ 
37:   if  $D \neq 0$  and  $T_{\max}$  increased then
38:     Pad all existing rows in  $D$  with NaN to maintain uniform dimension
39:   end if
40:   Pad current row to length  $1 + QT_{\max}$  with NaN values
41:   Append  $\tilde{x}_i$  to dataset  $D$ 
42: end for
43: Save  $D$  to disk as CSV file
44: return  $D$ 

```

For the sake of consistent length, all the samples are padded with NaN and stacked vertically into the dataset $\mathcal{D}$, which is then saved as a CSV file to be used for fault diagnosis. The dataset included a total of 43,365 samples, of which 30,355 samples (70%) were allocated for model development, and 13,010 samples (30%) were reserved as an independent test set. Within the development set, 80% of the samples were used for training, while the remaining 20% were utilized for validation.

Fig. 3 shows the waveforms used for the generation of the datasets in time-series format, illustrating various dynamic operating conditions such as system oscillations, equipment saturation, and nonlinear dynamics in the power grid system. These datasets simulate complex behaviors such as system oscillations and non-linearities, which affect the fault signatures under realistic operating scenarios. These datasets help

to validate the diagnostic tool by subjecting it to various fault conditions, thus improving the potential and ability of the tool to diagnose faults comprehensively.

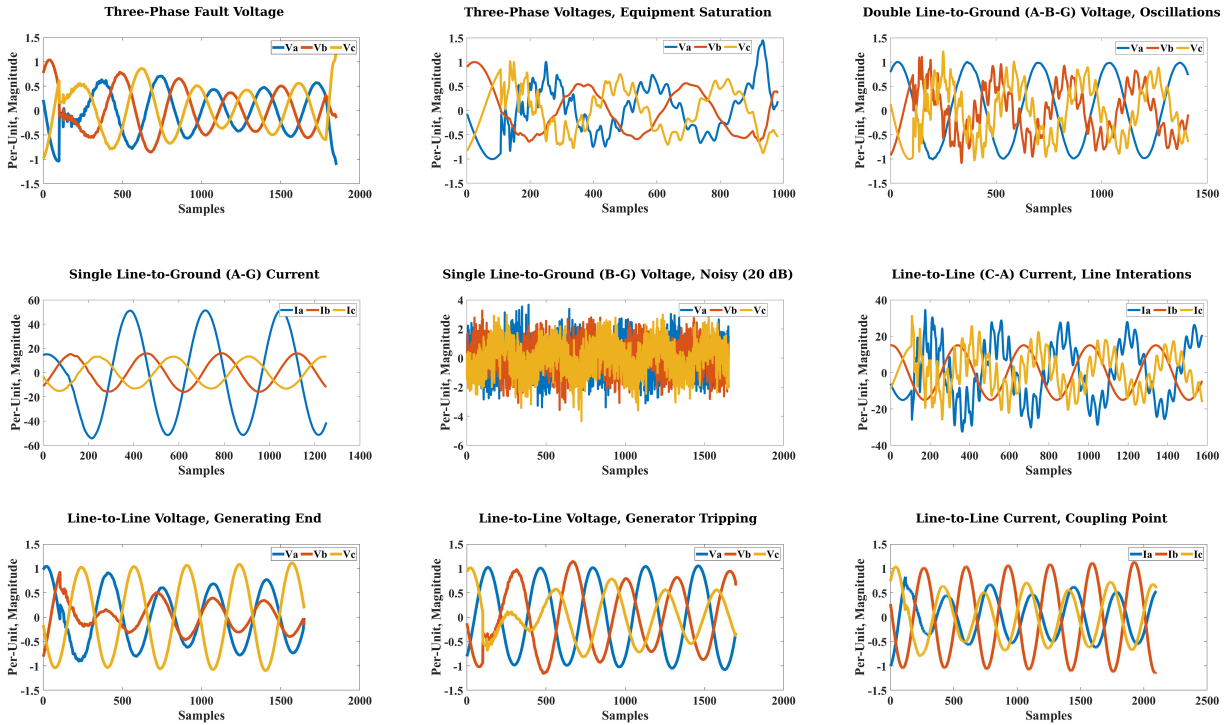


Figure 3: Representative three-phase voltage and current waveforms illustrating diverse fault conditions, including oscillations, saturation, noise effects, and line interactions in power system dynamics.

3 Results and Discussions

3.1 Accuracy and Efficiency Analysis

The performance analysis in Fig. 4a indicates that all three signal configurations have fault classification accuracy above 99%. Among the three, the voltage-current scheme has the highest performance with 99.91% test accuracy and the least standard deviation of $\pm 0.06\%$. It, however, has the highest training time of 1107 s and memory usage of 3330 MB among the three signal configurations. The voltage-only model has 99.72% accuracy while reducing the training time by 38.6% to 680 s and memory usage by 61.8% to 1272 MB, making it suitable for resource-constrained environments. The current-only model has 99.90% accuracy and performs best in real-time applications with the fastest inference speed of 2995 samples per second and quickest prediction time of 4.34 s, improving the processing throughput by 22.6%.

As can be seen in Fig. 4b, these methods have clear trade-offs between accuracy and computational efficiency. The voltage-current model is recommended for critical fault diagnosis systems where maximum accuracy is required; the voltage-only model is the best choice for edge computing scenarios where resources are limited; while the current-only model is the best option for high-frequency monitoring where real-time response is required. As can be seen, the viability of the three modalities of the signal is verified in the above analysis.

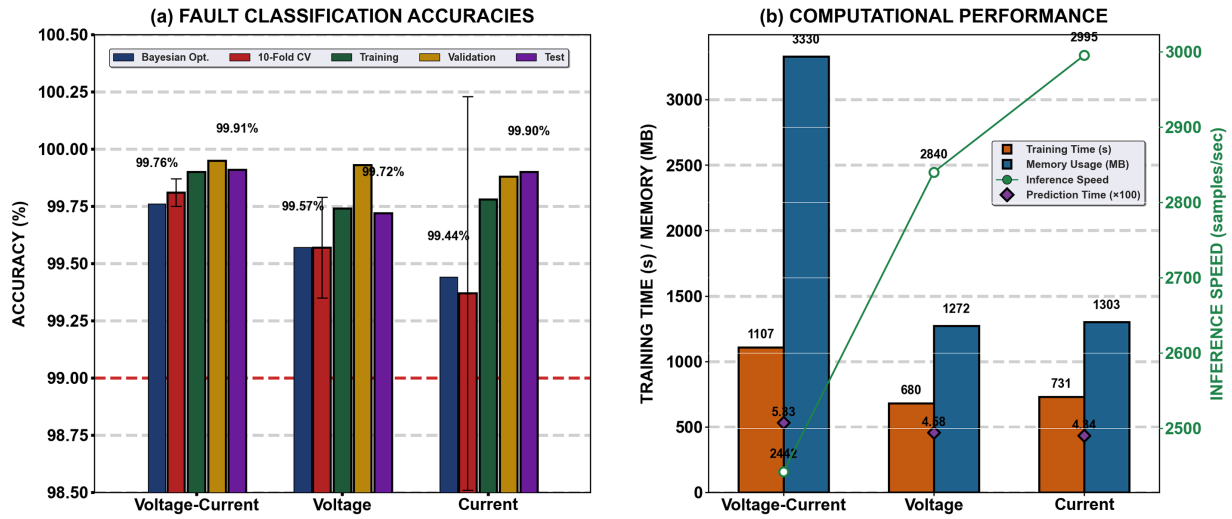


Figure 4: Performance comparison of signal types for UHV fault classification. (a) Accuracy metrics showing all models >99%, with voltage-current achieving highest test accuracy (99.91%) and best consistency ($\sigma = 0.06\%$). (b) Computational trade-offs: voltage-only reduces training time by 38.6% and memory by 61.8%; current-only delivers fastest inference (2995 samples/s).

3.2 Bayesian Optimization and Model Learning

The Bayesian optimization process achieved a best cross-validation accuracy of 99.76% at the twelfth iteration for the VI signals, as shown in Fig. 5a. The corresponding configuration used a learning rate of 0.002611, dropout rate of 0.182, batch size of 19, and weight decay of 8.65×10^{-6} . Across the evaluated configurations, the mean and median accuracies were 99.28% and 99.35%, respectively, while the standard deviation was 0.36 percentage points. The median being slightly higher than the mean suggests a modest tendency toward lower-valued observations, although this alone does not establish the shape of the accuracy distribution. The observed accuracy range was 1.63 percentage points, from 98.13% to 99.76%, indicating that most evaluated configurations achieved high classification accuracy despite variations in the selected hyperparameters. These results suggest that Bayesian optimization effectively explored the hyperparameter space and identified a high-performing configuration; however, the observed variation also indicates that hyperparameter selection can affect performance and should therefore be validated on independent test data before deployment.

The proposed architecture contains 913,314 trainable parameters, providing sufficient capacity to learn complex patterns associated with the UHV power system configuration. As shown in Fig. 5b, the validation performance improved up to epoch 27, where the validation loss reached 0.0031 and validation accuracy reached 99.95%. The training loss decreased from 0.4698 to 0.0093, corresponding to a 98% reduction, while the model achieved a training accuracy of 99.72% at the final epoch. After epoch 27, the validation loss increased slightly to 0.0038, and validation accuracy decreased marginally to 99.92% at epoch 30. This small degradation indicates a minor decline in validation performance during the final training epochs and suggests the onset of overfitting, although the final validation accuracy remained high. The training-validation accuracy difference at the final epoch was -0.20 percentage points, indicating that the validation accuracy remained slightly higher than the training accuracy. Overall, the model maintained high validation performance throughout training, with losses below 0.01 after epoch 10; however, the observed degradation after epoch 27 indicates that epoch 27 may provide a preferable stopping point based on validation performance.

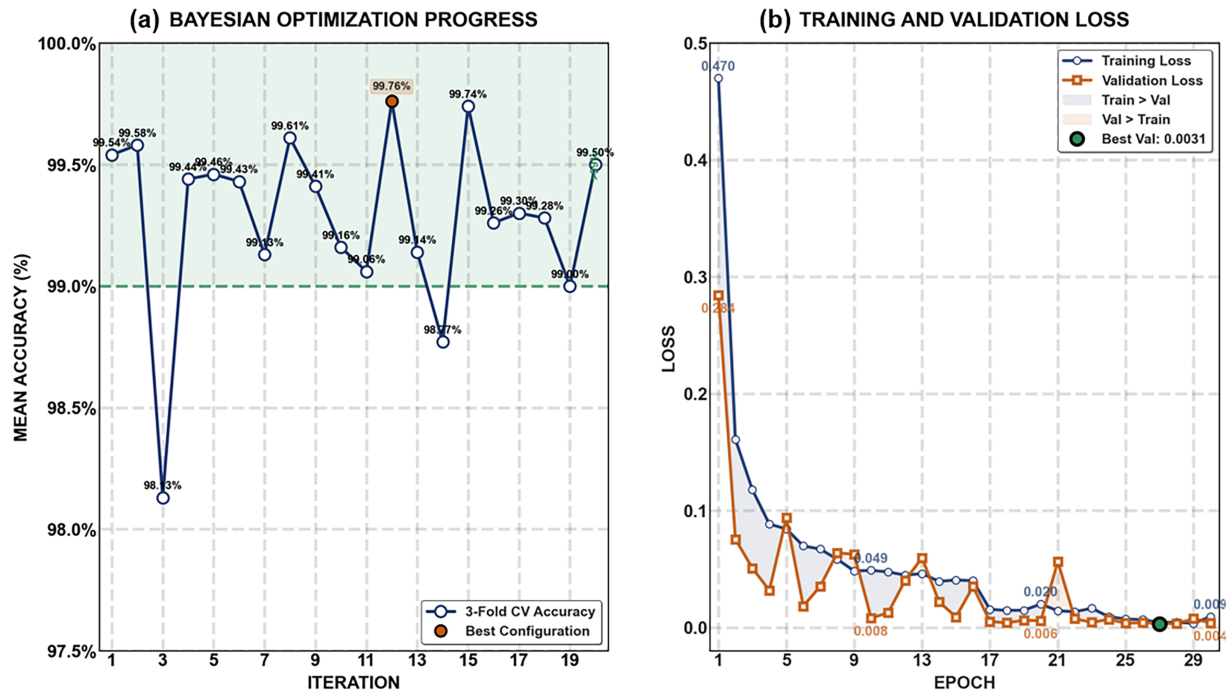


Figure 5: Optimization and training performance. (a) Bayesian optimization over twenty iterations achieves 99.76% peak accuracy at iteration twelve (orange), with most configurations exceeding 99% (green region). (b) Training and validation loss curves over thirty epochs show the best validation loss of 0.0031 at epoch twenty-seven (green point).

3.3 Feature Representation and Discriminative Capability Analysis

The outcomes of feature analysis also show the representation learning process of the proposed neural architecture. As illustrated in Fig. 6, the ratio of sparsity increased gradually from the shallow to deep layers of the network, reaching over 0.84 in conv3, global_ctx1, global_ctx3 and fc1. It can be seen from the observation that the deeper layers were able to learn more selective feature representations as a result of removing redundancy from the features while keeping the discriminative fault information intact. At the same time, the positive weight ratios stayed relatively well-balanced at about 0.5 for most layers, implying stable feature learning without any notable bias towards positive or negative kernel responses.

The kernel energy analysis showed that the L2 norms increased noticeably in the deeper layers, especially in the fully connected part of the network, where fc1, fc2, and fc3 reached 87.24, 52.78, and 27.89, respectively. This suggests that these higher-level layers captured more meaningful fault-related patterns and played a stronger role in making the final decision. On the other hand, the global context layers had much smaller L2 norms, which indicates that they mainly acted as lightweight attention modules for refining features rather than directly extracting the main representations.

The matrix of similarity of features indicates that the classifier is able to acquire very coherent features within each fault class while at the same time maintaining a high level of class separability. All diagonal elements of this matrix are greater than 0.86; in some classes, they exceed 0.98, which means that the features within each type of fault form clusters which are very coherent. Simultaneously, the low level of similarity off the diagonal, particularly between uncorrelated fault groups, indicates that the classifier is able to discriminate between classes. This is also evident from the results of class separability, including the Fisher separability index of 2.33 and the high level of between-class scatter relative to within-class scatter.

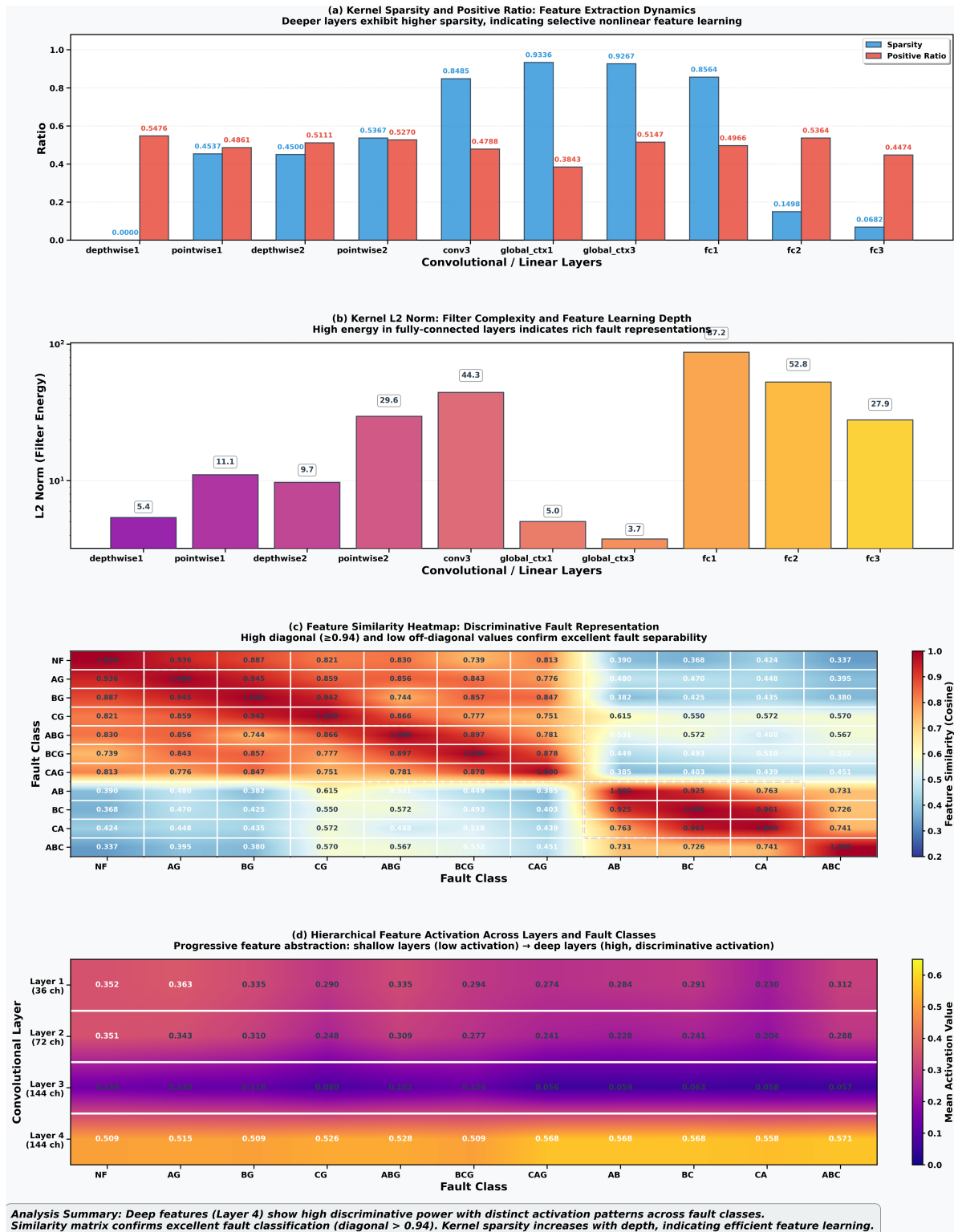


Figure 6: Analysis of learned representations in the proposed 1D CNN model: (a) kernel sparsity and positive-weight ratios, (b) kernel L2 norms, (c) inter-class feature similarity heatmap, and (d) layer-wise feature activations, demonstrating progressive feature abstraction and effective fault-class discrimination.

The hierarchical activation analysis shows how the model gradually learns more meaningful fault features as the data moves deeper through the network. In the first two convolutional layers, the activations are moderate, mainly capturing basic local patterns in the signals. The third layer becomes more selective, with lower overall activation but higher sparsity, meaning it focuses only on the most important features. By the final feature layer, the activations become stronger and more consistent—around 0.51 to 0.57 across different fault classes—showing that the network has successfully converted raw signal information into compact and clearly distinguishable fault representations. The clear differences in activation patterns between fault types further confirm that the model is effectively learning class-specific characteristics needed for accurate diagnosis.

3.4 Performance Generalization and Overfitting Characteristics

Fig. 7a showed evaluation on held-out testing data (30% of total samples) demonstrates exceptional performance with 99.9078% accuracy, correctly classifying 12,998 out of 13,010 instances with only twelve misclassifications across eleven fault categories. This further confirms the fact that the model learns discriminative features and is able to effectively transfer them to new examples. As depicted in Fig. 7b, the F1-score for Classes 1, 10, and 11 is perfect with a score of 1.0000. For Classes 2, 4, 7, and 9, the F1-score is above 0.9990, and the errors are spread out over a few categories. This further confirms the fact that the chosen hyperparameters result in a model with a proper bias-variance trade-off, which is essential in real-world applications.

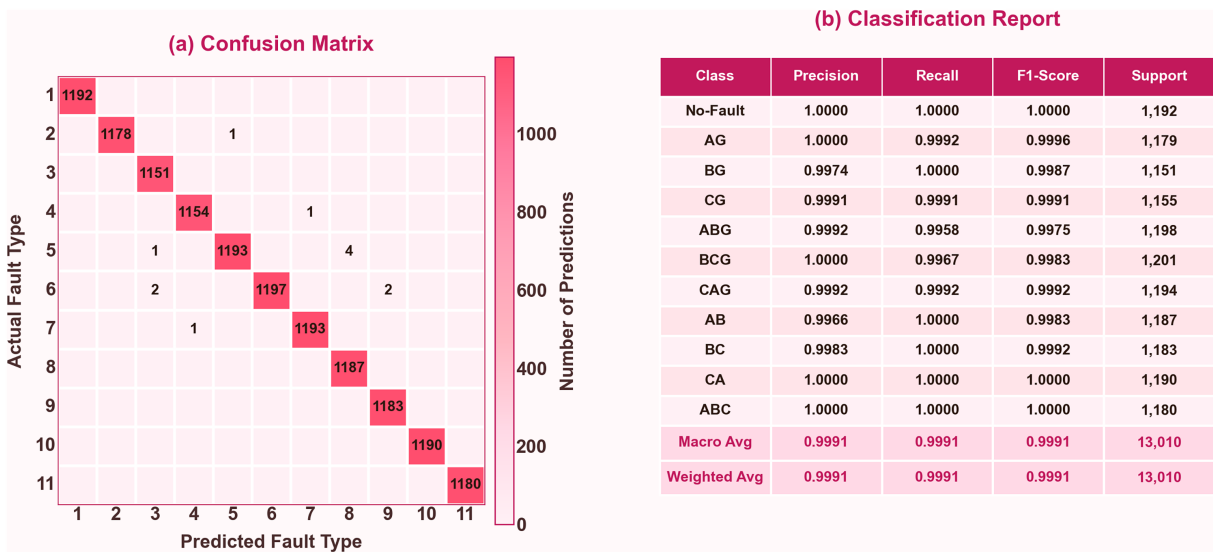


Figure 7: Comprehensive evaluation of the best-performing fault classification model on held-out test data (30% of total samples).

3.5 Noise Analysis

To assess the efficacy of the 1D CNN, noise analysis within a range of 10–50 dB is conducted with VI signals, which confirms the methodology's noise immunity under noisy conditions. The corruption of fault signals with WGN is an estimate of the methodology's robustness, as power system sensors are found to show Gaussian-type noise distortion in real-time operating conditions. As shown in Fig. 8a, it is clear that there is strong resilience to all types of noise. The performance of the classifier is found to show significant patterns under high noise conditions. At 10 dB, it is found to show 95.97% cross-validation and 97.03% test accuracy,

while at 20 dB, it is found to show 98.27% cross-validation and 98.28% test accuracy. Although there is a small decline in performance at 30 dB, with 97.56% test accuracy, it is clear from the results obtained at 40 and 50 dB, with 99.57% and 99.25% test accuracy, respectively, that there is effective handling of all types of signal-to-noise ratios.

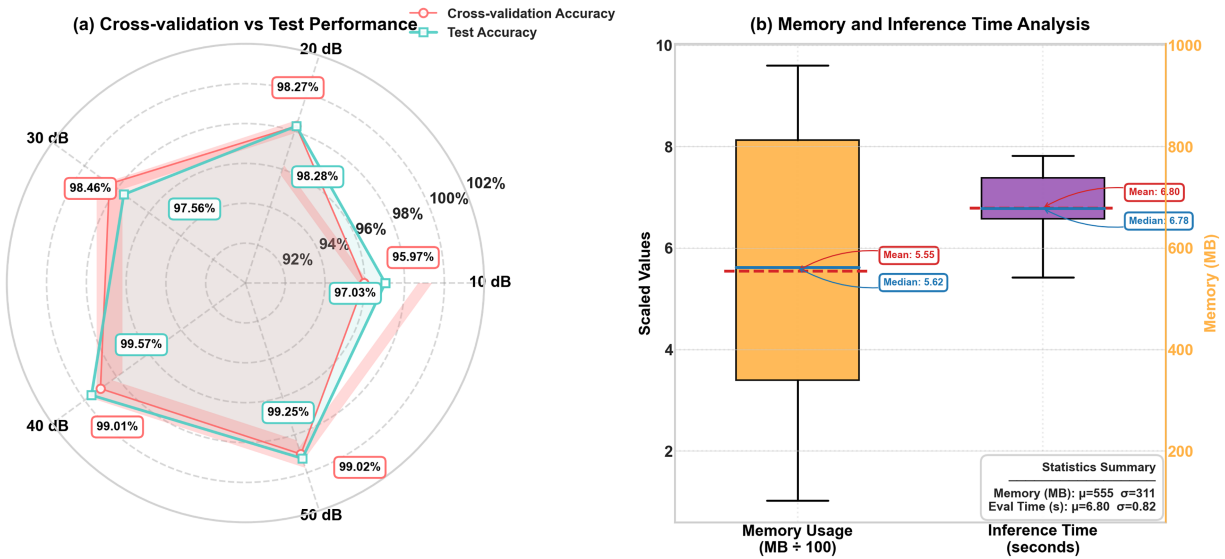


Figure 8: (a) Radar plot showing cross-validation and test accuracy across 10–50 dB noise levels, with error bands indicating standard deviation. (b) Box plot showing memory usage and inference time distribution with mean (red dashed) and median (blue solid) alignment indicating stable performance.

In Fig. 8b, memory usage varies between 102.11 MB for 50 dB and 959.97 MB for 40 dB. The average memory usage is 555 MB, while the median is 562 MB. This shows symmetry in memory usage for all the noise levels. The average inference time is 6.80 s, while the median is 6.78 s. The standard deviation for inference time is 0.82 s. This shows consistent performance in real-time processing. The classifier shows high immunity to all levels of noise. This eliminates the need for computationally expensive methods for noise cancellation or heavy preprocessing as reported in [1,24]. The accuracy for all levels of noise is above 97.03%, even for the highest noise level of 10 dB.

However, the similarity of mean and median for memory usage and inference time does not prove a normal distribution of the data but shows similar central tendencies among the analyzed conditions. Variations of memory usage and inference time show how the computational needs change in various noise levels. However, the classifier has demonstrated very good accuracy under different noise levels, that show consistency of the fault diagnosis process. Thus, these findings serve as proof of the robust fault pattern recognition by the classifier under the analyzed conditions. Comprehensive assessment across a broader range of operating conditions, noise levels, system configurations, and real-world scenarios would be required to further validate generalization, scalability, and deployment performance in critical-infrastructure environments.

The Noise Immunity Index (NII), defined in (8), provides a standardized quantitative measure of model robustness against noise by considering the classification performance across the entire investigated SNR range rather than only the extreme operating conditions. A lower NII indicates greater immunity to noise and more stable performance under varying operating environments.

$$NII = \frac{1}{M} \sum_{i=1}^M \frac{|V_{n,\max} - V_{n,i}|}{V_{n,\max}} \quad (8)$$

where M denotes the total number of investigated SNR levels, $V_{n,\max}$ represents the classification accuracy corresponding to the highest SNR level (lowest noise condition), and $V_{n,i}$ denotes the classification accuracy at the i -th SNR level. The index therefore quantifies the average normalized degradation in performance over the complete noise range. A value of NII approaching zero indicates that the model maintains nearly constant accuracy despite increasing noise levels, demonstrating strong robustness and practical applicability in real-world power system environments. Using the cross-validation accuracies of [95.97%, 98.27%, 98.46%, 99.01%, 99.02%] at 10, 20, 30, 40, and 50 dB SNR, respectively, with $V_{n,\max} = 0.9902$, the NII is calculated as

$$\begin{aligned} NII &= \frac{1}{5} \left[\frac{|0.9902 - 0.9597|}{0.9902} + \frac{|0.9902 - 0.9827|}{0.9902} + \frac{|0.9902 - 0.9846|}{0.9902} + \frac{|0.9902 - 0.9901|}{0.9902} + \frac{|0.9902 - 0.9902|}{0.9902} \right] \\ NII &= \frac{1}{5} \left[\frac{0.0305}{0.9902} + \frac{0.0075}{0.9902} + \frac{0.0056}{0.9902} + \frac{0.0001}{0.9902} + 0 \right] \\ &= \frac{1}{5} (0.03080186 + 0.00757423 + 0.00565542 + 0.00010099 + 0) \\ &= \frac{0.04413250}{5} \\ &= 0.00882650. \end{aligned}$$

The NII is calculated as 0.008826, corresponding to approximately 0.88% average normalized degradation. This low NII value indicates that the model maintained relatively consistent classification accuracy across the investigated SNR levels from 10 to 50 dB.

3.6 Sampling Analysis

This methodology is tested with voltage-only (V), current-only (I), and combined VI signals at various frequencies (1–20 kHz). The performance of the signals has been presented in Fig. 9, where the performance of the VI signals has been the most consistent. The performance of the signals has reached $99.81 \pm 0.06\%$ cross-validation accuracy and 99.91% test accuracy at 20 kHz. The performance of the voltage-only signals has reached 99.92% test accuracy at 5 kHz with $99.69 \pm 0.15\%$ cross-validation accuracy. The performance of the current-only signals has reached 99.90% test accuracy at 20 kHz with $99.37 \pm 0.86\%$ cross-validation accuracy.

As shown in Fig. 9c, for current-only signals, the training time is the least at 400.3 s compared to VI signals at 623.6 s and voltage-only signals at 545.6 s. Current-only signals have a reduction of 60% compared to VI signals. For inference speed, voltage-only signals have the highest at 14,213 samples per second, followed by current-only signals at 14,124 samples per second, and then VI signals at 11,228 samples per second. Current-only signals have the least memory footprint, while voltage-only and VI signals are moderate. The optimal accuracy-efficiency balance is between 5 and 15 kHz. As shown in Fig. 9d, for VI signals, there is consistent performance in terms of accuracy and efficiency for fault classification.

The proposed methodology is flexible in terms of sampling rates used in fault signals. It can reduce the computational load of data processing devices. The proposed methodology overcomes the limitations of the existing methodologies in terms of the range of sampling frequency used in the signals of the complex power system. The accuracy of the learning algorithms is generally decreased with the downsampling process. However, in the proposed methodology, it is observed that the effects of the downsampling process between the frequency range of 5–20 kHz on the VI signals are less significant. The proposed methodology

is satisfactory even at the frequency of 1 kHz. The reason is that the recording of the fault signals is obtained at different locations in the power system in the proposed methodology, which provides more fault information. Despite the loss of information in the significant downsampling process, the proposed methodology is more computationally efficient in the fault diagnosis of the power system compared with the existing methodologies that require the sampling frequency of 10–100 MHz.

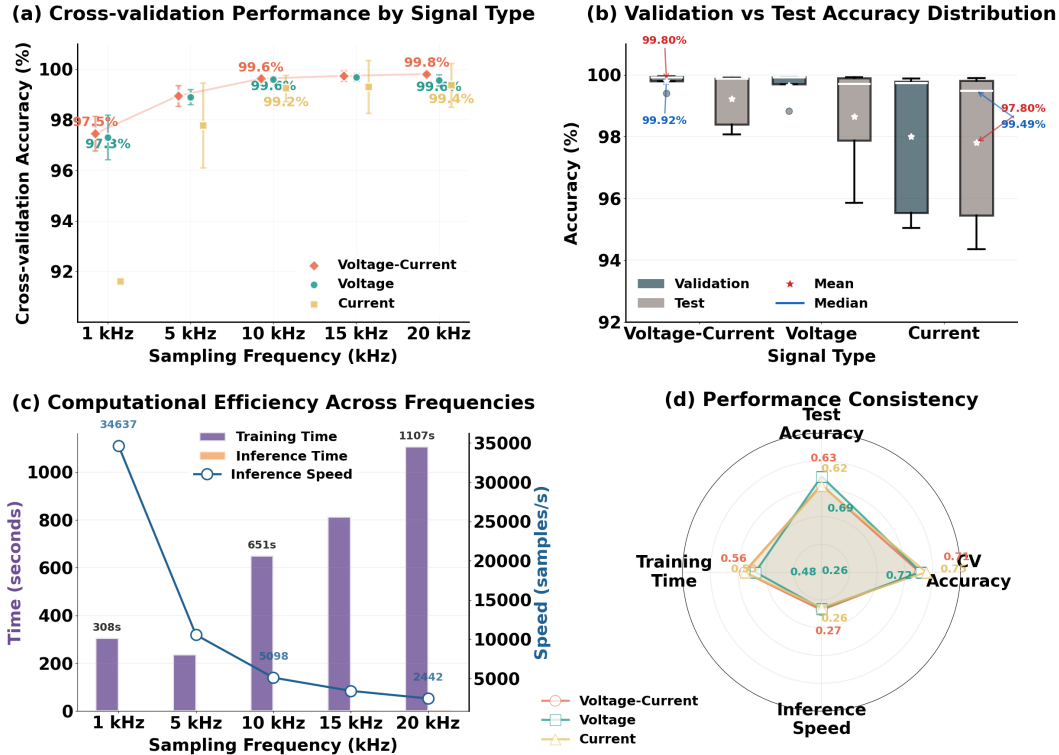


Figure 9: Performance metrics vs. sampling frequency: (a) Cross-validation accuracy, (b) Accuracy distribution, (c) Computational efficiency, (d) Normalized performance radar. V-I signals demonstrate optimal balance across all metrics.

The Sampling Index (SI), defined in (9), evaluates the sensitivity of the proposed approach to variations in sampling frequency by considering the entire investigated sampling range.

$$SI = \frac{1}{K} \sum_{j=1}^K \frac{|V_{s,max} - V_{s,j}|}{V_{s,max}} \quad (9)$$

where K denotes the total number of investigated sampling frequencies, $V_{s,max}$ is the classification accuracy corresponding to the highest sampling frequency, and $V_{s,j}$ is the classification accuracy at the j^{th} sampling frequency. The smaller SI values show that the suggested technique is stable with respect to various sampling frequencies, meaning that there is relatively low dependency on the data acquisition environment. The obtained SI value is 0.006933, which is less than 1%. This value is calculated using cross-validation accuracy at different sampling frequencies and reflects the stability of the proposed approach to signal downsampling. Although the diagnostic performance of the system varies depending on the data acquisition rate, it still shows stable performance regardless of this factor. This means that the proposed technique has stable performance at sampling frequencies in the 2–4 kHz range, which is considered the most optimal [27]. These findings confirm the method's robustness under sampling conditions.

3.7 The Perfect Recipe: Optimized Hyperparameters from Bayesian Optimization

As presented in Table 3, the optimized hyperparameters obtained via the application of the Bayesian optimization approach show the flexibility of the approach in adapting to the conditions of the analysis scenario. For the analysis scenario of the preliminary condition, the optimized parameters are determined to be batch size = 19, learning rate = 0.002611, dropout rate = 0.182, weight decay = 8.65×10^{-6} , and epochs = 30. These parameters serve as an excellent foundation for the training of the model. When the conditions are noisy at 20 dB, the parameters are optimized to be batch size = 79, learning rate = 0.001027, dropout rate = 0.488, weight decay = 5.74×10^{-4} , and epochs = 30. These parameters are optimized to be more regularized in the presence of noise in the signal.

Table 3: Optimized hyperparameters from Bayesian optimization for various analysis scenarios.

Hyperparameters	Preliminary Analysis	Noise Analysis (20 dB)	Sampling Analysis (10 kHz)
Batch size	19	79	18
Learning rate	0.002611	0.001027	0.003642
Dropout rate	0.182	0.488	0.177
Weight decay	8.65×10^{-6}	5.74×10^{-4}	1.50×10^{-5}
Number of epochs	30	30	30

Note: Bayesian optimization was employed to identify optimal hyperparameter configurations for each analysis scenario, ensuring peak model performance under varying conditions.

For sampling analysis with a sampling rate of 10 kHz, the hyperparameters are similar to those obtained by preliminary analysis, i.e., batch size is set to 18, the learning rate is set to 0.003642, the dropout rate is set to 0.177, weight decay is set to 1.50×10^{-5} , and the number of epochs is set to 30. This similarity implies that the model architecture is effective in handling different sampling rates without significant changes to the hyperparameters. Moreover, the consistent number of epochs across all scenarios implies that the model converges well within this number of epochs. Similarly, the dropout rate changes from 0.18 to 0.49 with the introduction of noise. This implies that the Bayesian optimization methodology recognizes the need to impose stronger regularization with the degraded signal. Similarly, the weight decay changes by almost two orders of magnitude with the introduction of noise. This is consistent with the observation that stronger regularization is required with the degraded signal. The changes to the learning rate by the optimization methodology reflect the trade-off between convergence speed and stability.

3.8 Comparison to Existing Work and Future Scope

Table 4 highlights that the proposed 1D CNN framework achieves superior accuracy even when applied to a highly complex UHV power grid configuration, which involves a 735 kV, 32-bus system with extensive compensation and multiple transmission lines. Although most of the currently available solutions may demonstrate high performance in some operational cases, those solutions require complex preprocessing, including the transformation of faults into image representations, thus increasing computational costs and distancing from the initial signal form. In contrast to that, the proposed solution allows classifying faults using the signal level without any additional processing and expert knowledge required.

Table 4: Performance comparison across fault diagnosis methods is constrained by variations in system configurations, fault scenarios, and evaluation protocols. The proposed 1D CNN achieves superior accuracy (99.91%) with a competitive memory footprint (1271.57–3329.95 MB) and faster inference (0.33–0.41 ms) at flexible sampling rates (1–20 kHz) in significantly complex power system configuration, outperforming existing approaches.

Ref.	Method	Scenario	Feature Extrac- tion./Selection	Sys. Config.	Accur. (%)	Data Signa.	Ser. Comp.	Data. (Size)	Train./Infer. Time	Noise Immun. (dB)	Learn. Type	Memory Usage MB	Samp. Analy. kHz	Preproc.
[36]	ANN- RFC	FDC	Feedforward +ReLU	Multi- Generator system	99.80	VI	No	10,000 [N]	2–3 h/15 ms	97.2 (20)	Hybrid	–	–	Yes
[40]	STE	FDCL	CFS	735, 32-bus, PG	98.30	17-[V, I]	Yes	30,950 [N]	2712.8 s	92.37 (20)	SL	1375.1	1–20	Yes
[39]	CNN- BFO	PV-FD	BFO	Kaggle PV Images	98.75	I	–	896 [I]	–	–	Hybrid	–	–	Yes
[12]	OSH- ENS	ADF	No	IEEE 14, 39	97.93	V, F, A	No	210,000 [N]	/21 ms	–	SL	–	20	No
[30]	MODWT- XGB	MG-FDC	MODWT	IEC-1457 MG, IEEE 13	99.57	2-[I]	No	1585 [N]	43.76 ms	98.31 (25)	SL	–	60 Hz	Yes
[29]	RF	FDC-PG	FSF	735 kV, 3-Bus	100	[V, F]	Yes	1100 [N]	1.2 cc	99.93 (15)	SL	–	5–20	Yes
[24]	SA- MNV3	FDC	CWT	735 kV, 3-Bus	99.90	1-[V, I]	Yes	20,735 [RGB]	588 s/19.8 ms	99.40 (15–35)	SL	–	5–20	Yes
[42]	LSTM	FDC-EHV	–	WSCC 9	99.45	[V, I]	No	27,000 [N]	8.33 ms	Yes (25–45)	SL	–	12	Yes
[1]	SAT- WSCN	FD	GAF	WSCC 9	99.81	1-[V, I]	Yes	38,115 [I]	387.3 s	99.70 (20)	SL	–	20	Yes
[27]	SAT- CNN	FDC	GAF-DWT	EHV	99.58	1-[V, I]	No	32,670 [I]	–	91.26 (10)	SL	–	2–20	Yes
[32]	CNSF	FDC	DWT	500 kV, EHV	99.72	2-[V, I]	No	38,115 [I]	0.52–0.64 cs	>99 (20)	UL	–	20	Yes
[17]	1D CNN	FCL	Automatic	Radial EPS	>99	1-[V, I]	No	13,000 [N]	347.75 s	–	SL	–	2	No
[28]	SDAE	FL	SDAE	RTDS	99.07	TW-[I]	No	400 [N]	–	96.73 (20)	SL	–	–	Yes
[43]	CISS- SVM	FDC	WPT	PSCAD /EMTP	98.82	1-[V, I]	Yes	[N]	–	82.00 (–)	SL	–	–	Yes
[22]	RFE- XGB	FDC	RFE	PTL	94.25	1-[V, I]	No	2040 [N]	–	–	SL	–	2	Yes

(Continued)

Table 4 (continued)

Ref.	Method	Scenario	Feature Extrac- tion./Selection	Sys. Config.	Accur. (%)	Data Signa.	Ser. Comp.	Data. (Size)	Train./Infer. Time	Noise Immun. (dB)	Learn. Type	Memory Usage MB	Samp. Analy. kHz	Preproc.
[14]	F-DTW- CNN	FC	STFT	PSCAD /EMTDC, RTDS	99.37	[V, I]	No	10,766 [I]	-	-	SL	-	-	Yes
[8]	TW-ML	FCL	SWT-MM	IEEE 34	99.43	[V, I]	No	15,282 [N]	-	93.7 (45)	SL	-	10 MHz	Yes
[20]	RF-GBC	IDS-CTM	RFE	PG [ORNL]	93.97	[V, I, A]	-	102,890 [N]	-	-	SL	-	-	Yes
[7]	DRNN	FDC	Numerical	Two-Area Four- Machine 11-Bus	96.71	VI	-	2562 [N]	323 s	-	SL	-	60 Hz	Yes
[18]	CNN- LSTM	FCL	Spatial-Temporal	IEEE 9	99.22	VI	No	51,840 [N]	-	-	SL	-	-	No
[16]	CNN- LSTM	FCL	Spatial-Temporal	IEEE 9	99.82	3-[V, I]	No	243,000 [N]	10,920 s	-	SL	-	-	No
[13]	CNN- GTO	APS	CNN-GTO	FREEDM, Microgrid	99.37	[V, I]	-	224,288 [N]	1.3 ms	98.14 (20)	SL	-	20	Yes
Proposed	1D CNN	FC	Automatic	735 kV, 32-Bus, PG	99.91	15-VI	Yes	43,365 [N]	237.94- 110736 s/0.33389- 0.409467 ms	98.42 (20)	SL	1271.57- 3329.95	1-20	No

Note: **Abbreviations:** PG: power grid; MG: microgrid; N: numerical; I: images; IDS: Intrusion Detection System; FC: Fault Classification; FD: Fault Detection or Diagnosis; FCL: Fault Classification and Location; APS: Adaptive Protection Scheme; SL: Supervised Learning; UL: Unsupervised Learning; MODWT: Maximal Overlap Discrete Wavelet Transform; ANFIS: Adaptive Neuro-Fuzzy Inference System; GAF: Gramian Angular Field; DWT: Discrete Wavelet Transform; GWO: Grey Wolf Optimizer; WELM: Weighted Extreme Learning Machine; F-DTW-CNN: Fast Dynamic Time Warping CNN; MM: Mathematical Morphology; RNN: Recurrent Neural Network; GCN: Graph Convolutional Network; MVE: Majority Vote Ensemble; DRNN: Deep Recurrent Neural Network; GTO: Gorilla Troops Optimizer; WSCC: Western Systems Coordinating Council; RTDS: Real-Time Digital Simulator; FREEDM: Future Renewable Electric Energy Delivery and Management; ORNL: Oak Ridge National Laboratory; VI: Voltage and Current; V,F,A: Voltage, Frequency, Angle; TW: Traveling Wave; PMU: Phasor Measurement Unit.

To ensure an unbiased comparison across diverse studies, the framework introduces uniform indices for sampling and noise immunity, allowing performance evaluation under varying data complexities and operating conditions. The results show that the proposed method not only achieves excellent accuracy but also maintains outstanding robustness under noisy environments and downsampling scenarios. Furthermore, it provides a comprehensive analysis across multiple signal types, which is particularly valuable in data-scarce conditions. Unlike most prior works, this study also reports memory usage, revealing that while VI signals demand higher resources, individual voltage or current signals achieve comparable efficiency to existing methods, making the approach adaptable depending on resource constraints. Importantly, the framework challenges fault classification in UHV systems by incorporating a system-wide analysis across all transmission lines, thereby enhancing generalization and scalability compared with studies limited to specific lines.

By replicating the complexities of a physical UHV system in a simulation, the proposed methodology demonstrates its practical relevance and confirms its potential as a robust solution for fault diagnosis in large-scale power grids.

4 Conclusions

The proposed study demonstrates the effectiveness of a lightweight 1D CNN classifier for fault classification in UHV power systems, validated using a standard simulation setup operating at UHV level with series-compensated and distributed transmission lines designed by domain experts. The real-time Simulink model replicates the complexities of practical large-scale power systems, ensuring realistic evaluation. The 1D CNN effectively addresses dynamic and challenging operating conditions, delivering high accuracy and computational efficiency through end-to-end learning on raw time-series fault signals, which aligns with modern sensor recordings and eliminates the need for heavy preprocessing or transformations.

The automatic feature engineering capability of the model surpasses traditional methods reliant on sophisticated feature extraction and selection, which may fail to capture the intricate behaviors of series-compensated UHV systems. Bayesian optimization for hyperparameter tuning enhances the robustness and efficiency of model development. The data processing pipeline achieves high noise immunity and mitigates the adverse effects of downsampling, providing operational flexibility for diverse data processing devices.

The introduction of quantitative comparative metrics enables fair evaluation across fault diagnosis methodologies that differ in system configurations, data complexity, and operating conditions. While the method effectively overcomes many limitations, further work is needed to optimize noise cancellation using advanced lightweight filtering and improve memory utilization when handling a large number of synchronous recordings, a common scenario in practical systems. Given its robustness, the proposed approach has potential extensions to intrusion detection and fault localization in power grids.

In summary, this rigorously evaluated methodology establishes a proof of concept for practical, scalable fault diagnosis in large-scale UHV power systems, combining accuracy, efficiency, and adaptability to real-world conditions.

Acknowledgement: The authors extend their appreciation to the Research, Development, and Innovation Authority (RDIA), Saudi Arabia, for funding this work through grant number (13385-Tabuk-2023-UT-R-3-1-SE).

Funding Statement: This article is derived from a research grant funded by the Research, Development, and Innovation Authority (RDIA)—Kingdom of Saudi Arabia—with grant number (13385-Tabuk-2023-UT-R-3-1-SE).

Author Contributions: Hani Albalawi: Conceptualization, methodology, and writing—review and editing; Rab Nawaz: Software, investigation, data curation, and writing—original draft; Abdul Wadood: Conceptualization, validation, supervision, and writing—review and editing; Anwar Ul Haq: Formal analysis, visualization, and validation; Shahbaz Khan: Software, investigation, and data analysis; Bakht Muhammad Khan: Methodology, resources, and writing—review and editing; Adel Mohammed Alatwi: Project administration, funding acquisition, and writing—review and editing. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the Corresponding Author.

Ethics Approval: Not applicable.

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

References

1. Fahim SR, Sarker SK, Muyeen SM, Sheikh MRI, Das SK, Simoes M. A robust self-attentive capsule network for fault diagnosis of series-compensated transmission line. *IEEE Trans Power Deliv.* 2021;36(6):3846–57. doi:10.1109/tpwr.2021.3049861.
2. Teta A, Korich B, Bakria D, Hadroug N, Rabehi A, Alsharef M, et al. Fault detection and diagnosis of grid-connected photovoltaic systems using energy valley optimizer based lightweight CNN and wavelet transform. *Sci Rep.* 2024;14(1):18907. doi:10.1038/s41598-024-69890-7.
3. Anwar T, Mu C, Yousaf MZ, Khan W, Khalid S, Hourani AO, et al. Robust fault detection and classification in power transmission lines via ensemble machine learning models. *Sci Rep.* 2025;15(1):2549. doi:10.1038/s41598-025-86554-2.
4. Sahoo PK, Mishra S, Swain SC. Classification of incipient fault in underground distribution networks using extreme gradient boost classifier. In: *Proceedings of the 2025 International Conference on Power Electronics and Energy (ICPEE)*; 2025 Dec 14–16; Bhubaneswar, India. New York, NY, USA: IEEE; 2025. p. 1–6.
5. Demirci M. Multi fault classification in electrical transmission lines using feature engineering based on autogluon framework. *Sustain Comput Inform Syst.* 2026;50(15):101310. doi:10.1016/j.suscom.2026.101310.
6. Fahim SR, Atat R, Kecici C, Takiddin A, Ismail M, Davis KR, et al. Graph neural network-based approach for detecting false data injection attacks on voltage stability. *IEEE Open J Power Energy.* 2025;12:12–23. doi:10.1109/oajpe.2024.3524268.
7. Belagoune S, Bali N, Bakdi A, Baadji B, Atif K. Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems. *Measurement.* 2021;177(3):109330. doi:10.1016/j.measurement.2021.109330.
8. Jimenez-Aparicio M, Wilches-Bernal F, Reno MJ. Local, single-ended, traveling-wave fault location on distribution systems using frequency and time-domain data. *IEEE Access.* 2023;11:74201–15. doi:10.1109/access.2023.3296737.
9. Series-compensated transmission system—MATLAB & Simulink—mathworks.com. [cited 2026 Mar 5]. Available from: <https://www.mathworks.com/help/releases/R2024b/sps/ug/power-grid-statcom.html>.
10. Yuvaraju V, Thangavel S, Golla M. Applications of artificial intelligence and PMU data: a robust framework for precision fault location in transmission lines. *IEEE Access.* 2024;12(6):136565–87. doi:10.1109/access.2024.3464088.
11. Mishra PK, Yadav A. Fault classification scheme for TCSC compensated transmission line using MODWPT-BTEC. *e-Prime Adv Electr Eng Electron Energy.* 2024;9:100729. doi:10.1016/j.prime.2024.100729.
12. Yavuz L, Soran A, Önen A, Li X, Muyeen S. Adaptive fault detection scheme using an optimized self-healing ensemble machine learning algorithm. *CSEE J Power Energy Syst.* 2021;8(4):1145–56. doi:10.17775/cseejpes.2020.03760.
13. Hatata AY, Essa MA, Sedhom BE. Adaptive protection scheme for FREEDM microgrid based on convolutional neural network and *Gorilla* troops optimization technique. *IEEE Access.* 2022;10:55583–601. doi:10.1109/access.2022.3177544.

14. Yang NC, Yang JM. Fault classification in distribution systems using deep learning with data preprocessing methods based on fast dynamic time warping and short-time Fourier transform. *IEEE Access*. 2023;11:63612–22. doi:10.1109/access.2023.3288852.
15. Fahim SR, Sarker MRI, Arifuzzaman M, Hosen MS, Sarker SK, Das SK. A novel approach to fault diagnosis of high voltage transmission line-a self attentive convolutional neural network model. In: *Proceedings of the 2020 IEEE Region 10 Symposium (TENSYP)*; 2020 Jun 5–7; Dhaka, Bangladesh. New York, NY, USA: IEEE; 2025. p. 1329–32.
16. Minh NQ, Khiem NT, Giang VH. Fault classification and localization in power transmission line based on machine learning and combined CNN-LSTM models. *Energy Rep*. 2024;12(1):5610–22. doi:10.1016/j.egy.2024.11.061.
17. Yoon DH, Yoon J. Deep learning-based method for the robust and efficient fault diagnosis in the electric power system. *IEEE Access*. 2022;10:44660–8. doi:10.1109/access.2022.3170685.
18. Alhanaf AS, Farsadi M, Balik HH. Fault detection and classification in ring power system with DG penetration using hybrid CNN-LSTM. *IEEE Access*. 2024;12:59953–75. doi:10.1109/access.2024.3394166.
19. Nawaz R, Wadood A, Khalid Mehmood K, Basit Ali Bukhari S, Albalawi H, Mohammed Alatwi A, et al. Gradient boosting feature selection for integrated fault diagnosis in series-compensated transmission lines. *IEEE Access*. 2025;13:63640–70. doi:10.1109/access.2025.3558796.
20. Zaman M, Upadhyay D, Lung CH. Validation of a machine learning-based IDS design framework using ORNL datasets for power system with SCADA. *IEEE Access*. 2023;11:118414–26. doi:10.1109/access.2023.3326751.
21. Nawaz R, Albalawi HA, Ali Bukhari SB, Khalid Mehmood K, Sajid M. Timeseries fault classification in power transmission lines by non-intrusive feature extraction and selection using supervised machine learning. *IEEE Access*. 2024;12(4):93426–49. doi:10.1109/access.2024.3423828.
22. Guo B, Yang B, Wang S, Shi W, Yang F, Wang D. Machine learning-based fault diagnosis and classification of three-phase transmission lines with RFE and domain knowledge. *Electr Power Syst Res*. 2025;247(2):111777. doi:10.1016/j.epsr.2025.111777.
23. Liu K, Li B, Jiang Q, Zhang Y, Liu T. Fault ride-through strategy for hybrid cascaded HVDC systems based on controllable LCC. *IEEE Trans Circuits Syst II*. 2026;73(1):88–92. doi:10.1109/tcsii.2025.3627209.
24. Xi Y, Zhang W, Zhou F, Tang X, Li Z, Zeng X, et al. Transmission line fault detection and classification based on SA-MobileNetV3. *Energy Rep*. 2023;9:955–68. doi:10.2139/ssrn.4226448.
25. Zheng X, Xu N, Trinh L, Wu D, Huang T, Sivaranjani S, et al. A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids. *Sci Data*. 2022;9(1):359. doi:10.1038/s41597-022-01455-7.
26. Xu W, Li J, Yang L, Yu Q. Faults locating of power distribution systems based on successive PSO-GA algorithm. *Sci Rep*. 2024;14(1):11259. doi:10.1038/s41598-024-61306-w.
27. Fahim SR, Sarker Y, Sarker SK, Sheikh MRI, Das SK. Self attention convolutional neural network with time series imaging based feature extraction for transmission line fault detection and classification. *Electr Power Syst Res*. 2020;187(5):106437. doi:10.1016/j.epsr.2020.106437.
28. Shang B, Luo G, Li M, Liu Y, Hei J. Transfer learning-based fault location with small datasets in VSC-HVDC. *Int J Electr Power Energy Syst*. 2023;151:109131. doi:10.1016/j.ijepes.2023.109131.
29. Najafzadeh M, Pouladi J, Daghigh A, Beiza J, Abedinzade T. Fault detection, classification and localization along the power grid line using optimized machine learning algorithms. *Int J Comput Intell Syst*. 2024;17(1):49. doi:10.1007/s44196-024-00434-7.
30. Patnaik B, Mishra M, Bansal RC, Jena RK. MODWT-XGBoost based smart energy solution for fault detection and classification in a smart microgrid. *Appl Energy*. 2021;285:116457. doi:10.1016/j.apenergy.2021.116457.
31. Almeida AR, Almeida OM, Junior BFS, Barreto LHSC, Barros AK. ICA feature extraction for the location and classification of faults in high-voltage transmission lines. *Electr Power Syst Res*. 2017;148:254–63. doi:10.1016/j.epsr.2017.03.030.
32. Fahim SR, Sarker SK, Muyeen SM, Das SK, Kamwa I. A deep learning based intelligent approach in detection and classification of transmission line faults. *Int J Electr Power Energy Syst*. 2021;133(4):107102. doi:10.1016/j.ijepes.2021.107102.

33. Chen K, Hu J, He J. Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder. *IEEE Trans Smart Grid*. 2016;9(3):1748–58. doi:10.1109/pesgm.2017.8273992.
34. Yousaf MZ, Singh AR, Khalid S, Bajaj M, Kumar BH, Zaitsev I. Bayesian-optimized LSTM-DWT approach for reliable fault detection in MMC-based HVDC systems. *Sci Rep*. 2024;14(1):17968. doi:10.1038/s41598-024-68985-5.
35. Abdelgayed TS, Morsi WG, Sidhu TS. A new harmony search approach for optimal wavelets applied to fault classification. *IEEE Trans Smart Grid*. 2016;9(2):521–9. doi:10.1109/tsg.2016.2555141.
36. Uzel H, Özüpak Y, Alpsalaz F, Aslan E. Optimized ANN-RF hybrid model with optuna for fault detection and classification in power transmission systems. *Sci Rep*. 2025;16(1):1495. doi:10.1038/s41598-025-31008-y.
37. Alam MS, Islam MR, Fan R, Alam Shazid MS, Hasan AS. Attention-guided multi-task learning for fault detection, classification, and localization in power transmission systems. *Energies*. 2025;18(24):6547. doi:10.3390/en18246547.
38. Biswas S, Nayak PK, Panigrahi BK, Pradhan G. An intelligent fault detection and classification technique based on variational mode decomposition-CNN for transmission lines installed with UPFC and wind farm. *Electr Power Syst Res*. 2023;223(3):109526. doi:10.1016/j.epsr.2023.109526.
39. Sabati A, Bayindir R, Rahebi J. Photovoltaic panels fault detection with convolutional neural network and bitterling fish optimization (BFO) algorithm. *Int J Comput Intell Syst*. 2025;18(1):239. doi:10.1007/s44196-025-00984-4.
40. Nawaz R, Qamar A, Wadood A, Albalawi H. Consensus-based feature selection and Bayesian-optimized stacking ensembles for comprehensive fault diagnosis in series-compensated ultra-high-voltage power systems. *Sci Rep*. 2026;16(1):18678. doi:10.1038/s41598-026-49195-7.
41. Ukwuoma CC, Cai D, Bamisile O, Chukwuebuka EJ, Favour E, Emmanuel GSA, et al. Power transmission system's fault location, detection, and classification: pay close attention to transmission nodes. *Int J Electr Power Energy Syst*. 2024;156(6):109771. doi:10.1016/j.ijepes.2023.109771.
42. Rafique F, Fu L, Mai R. End to end machine learning for fault detection and classification in power transmission lines. *Electr Power Syst Res*. 2021;199(10):107430. doi:10.1016/j.epsr.2021.107430.
43. Kumar VR, Jeyanthi PA. Fault classification and detection in transmission lines by hybrid algorithm associated support vector machine. *Trans Emerging Tel Tech*. 2025;36(3):e70034. doi:10.1002/ett.70034.