A Compact and Robust Framework for Multi-Condition Transient Pressure-Wave-Based Leakage Identification in District Heating Networks

Chang Chang , Xiangli Li * , Xin Jia , Lin Duanmu

Posted Date: 19 March 2026

doi: 10.20944/preprints202603.1448.v1

Keywords: district heating networks (DHNS); leakage detection; fault diagnosis; ReliefF; GAN-based data augmentation; multi-condition classification

 doi:10.20944/preprints202603.1448.v1

Article

A Compact and Robust Framework for Multi-Condition Transient Pressure-Wave-Based Leakage Identification in District Heating Networks

Chang Chang [1], Xiangli Li [1,*], Xin Jia [2] and Lin Duanmu [1]

[1] Institute of Building Environment and Facility Engineering, Dalian University of Technology, Dalian 116024, China

[2] College of Civil Engineering and Architecture, Dalian University, Dalian 116622, China

* Correspondence: lxl@dlut.edu.cn

Abstract

Leakage identification in district heating networks (DHNS) is challenging because leakage-induced transient pressure waves often overlap with pressure disturbances triggered by routine operations such as valve regulation, pump speed variation, and emergency shut-off. In addition, the scarcity of high-quality labeled leakage samples limits the robustness of data-driven models under small-sample conditions. To address these issues, this study proposes a compact and moderately interpretable framework for multi-condition identification from transient pressure-wave signals, integrating signal preprocessing, handcrafted statistical feature extraction, multiclass ReliefF-based feature selection, and class-wise generative adversarial network (GAN) augmentation in the selected feature space. A dataset containing four representative conditions, namely leakage, valve regulation, pump speed regulation, and emergency valve shut-off, was constructed using an integrated indoor district heating network testbed. After Hampel-based spike suppression and zero-phase Butterworth band-pass filtering within 0.5 to 300 Hz, time- and frequency-domain statistical features were extracted, and a compact subset was selected by multiclass ReliefF. A class-wise GAN was then used to augment the training set in feature space, while all evaluations were performed strictly on real samples. The results show that feature-space augmentation improves robustness and generalization under operational disturbances and noise. Using random forest as the representative classifier, Accuracy and Macro-F1 increased from 0.960 to 0.985, while leakage recall improved from 0.920 to 0.980. Further comparisons confirmed that the ReliefF-selected subset outperformed representative alternatives such as LASSO and mRMR. Overall, the proposed framework provides an effective solution for distinguishing leakage events from operational disturbances and offers practical support for online monitoring and intelligent operation of DHNS.

Keywords: district heating networks (DHNS); leakage detection; fault diagnosis; ReliefF; GAN-based data augmentation; multi-condition classification

1. Introduction

As a critical component of urban infrastructure, the safe and stable operation of district heating networks (DHNs) directly affects energy efficiency, public safety, and the quality of life of residents. With the continuous expansion of DHN scale, the increasing complexity of network configurations, and the rising frequency of operational adjustments, heat carrier losses, pressure imbalances, reduced transmission efficiency, and operational risks caused by pipeline leaks have become increasingly prominent. Previous studies have shown that faults such as leaks, abnormal heat losses, and pressure drops significantly impact system reliability and operational economy. Therefore, developing rapid, accurate, and online-applicable leak identification methods is of great engineering significance [1–3].

 doi:10.20944/preprints202603.1448.v1

In the field of DHN leak detection, existing methods can be broadly classified into external and internal diagnostic approaches. External methods, such as infrared thermography, acoustic sensing, and fiber-optic sensing, typically locate leaks by detecting changes in temperature, acoustic, or strain fields induced by the leakage event [4–6]. However, these techniques are often constrained by environmental conditions, burial methods, and on-site deployment limitations, making it difficult to achieve stable real-time online monitoring. In contrast, internal diagnostic methods rely on operational parameters like pressure and flow, rendering them more suitable for continuous monitoring and online deployment.

Among various internal diagnostic techniques, pressure-wave-based leak detection has gained considerable attention in recent years due to its fast response, online monitoring capability, and relative insensitivity to external environmental variations. These methods can be categorized into passive and active pressure-wave detection. The former utilizes negative pressure waves or acoustic emission signals spontaneously generated by the leak itself, analyzing their propagation characteristics for leak localization and identification [7–9]. The latter injects controlled pressure waves into the pipeline via an external excitation device and diagnoses leaks based on the reflection, attenuation, or spectral changes of the waves at the leak point [10,11].

Nevertheless, during actual operation, routine actions such as valve regulation, pump start-up/shutdown, or speed adjustments also induce significant transient pressure fluctuations whose signatures partially overlap with leak signals, posing a serious challenge to the accuracy and robustness of leak identification. Moreover, measurement noise can substantially affect pressure-wave assessment and leak localization accuracy [8–13]. In parallel, research on DHN fault diagnosis has evolved from physical modeling and simulation-based optimization to data-driven approaches. Existing studies have covered the detection and identification of typical faults such as water leakage, abnormal heat loss, and pressure loss, as well as leakage diagnosis and localization based on sensor measurements [1–3,14–16]. Overall, DHN fault diagnosis is gradually shifting from purely mechanism-based analysis toward integrated model-driven and data-driven paradigms. However, research on fine-grained multi-class identification of pressure-wave signals under complex operating conditions remains relatively scarce. Hence, stably extracting discriminative leak features from pressure-wave signals under multiple disturbances, strong noise, and complex scenarios has become a critical issue urgently needing resolution.

In practical industrial environments, acquiring leak samples is difficult, and the amount of high-quality labeled data available for model training is typically limited, which has become a major bottleneck restricting the performance improvement of machine-learning methods. To address this small-sample problem, researchers have explored transfer learning and generative modeling. Tang et al. proposed a transfer-learning framework for detecting small-aperture leaks in natural gas pipelines [17], while Shi et al. developed a small-sample diagnostic method for pipe burst detection in water distribution systems [18]. In the area of generative adversarial networks (GANs), Wang et al. alleviated the small-sample problem in oil and gas pipeline leakage detection by improving GAN-based modeling [19]; Rajabi et al. employed conditional deep convolutional GANs for leakage detection and localization, demonstrating the potential of generative approaches for sample augmentation and identification in pipeline networks [20]; Zhang et al. and Shang et al. further improved data generation quality and class adaptability in industrial pipeline fault diagnosis through hybrid generative networks and multi-generator switching mechanisms, respectively [21,22]. These studies indicate that generative approaches can mitigate data scarcity to some extent. However, most existing work focuses on raw signals or other industrial scenarios, while small-sample augmentation tailored to the statistical feature space of DHN pressure waves remains insufficiently explored.

Beyond sample augmentation, feature construction and selection are also crucial for improving identification performance under small-sample conditions. High-dimensional small-sample classification is prone to the curse of dimensionality and overfitting, prompting the development of various feature selection strategies [23]. For instance, mutual-information-based methods entail high computational costs [24]; sparse representation methods such as LASSO suffer from limited stability

 doi:10.20944/preprints202603.1448.v1

when the sample size is very small [25]; and embedded evaluations like those in random forests can be influenced by hyperparameters and may retain redundant information [26]. By contrast, ReliefF, as a multi-class extension of the Relief algorithm, updates feature weights by comparing the differences between nearest neighbors from the same class and different classes, thereby effectively capturing local discriminative information that contributes significantly to decision boundaries [27–29]. This algorithm offers high computational efficiency, decoupling from the downstream classifier, robustness to noise, and no requirement for feature independence, making it particularly suitable for high-dimensional feature screening in small-sample settings. In the leakage detection field, Liang et al. and related studies have verified the effectiveness of ReliefF-like methods for classifying acoustic emission signals of valve internal leakage and for cross-domain pipeline defect diagnosis [30,31].

In summary, although considerable progress has been made in DHN leak detection, pressure-wave analysis, small-sample learning, and feature selection, notable gaps remain in fine-grained multi-class identification of pressure-wave signals under complex operating conditions. These gaps include the confusion between operational disturbances and leak features, insufficient model generalization with limited samples, limited research on small-sample augmentation in the statistical feature space of DHN pressure waves, and a lack of systematic studies on constructing compact, robust, and interpretable feature subsets. To address these issues, this study develops a compact, robust, and moderately interpretable multi-condition identification framework tailored to pressure-wave signals from DHNs under complex operating conditions, where signal confusion and limited sample sizes hinder model generalization. The proposed framework integrates statistical feature extraction, multiclass ReliefF-based feature selection, and GAN-based feature-space augmentation to enable fine-grained multi-class identification that distinguishes leaks from typical operational disturbances.

2. Experimental System and Dataset

2.1. Experimental System

Leakage events in district heating networks (DHNs) are typically accompanied by transient pressure-wave propagation [7]. Compared with approaches that rely on steady state measurements, leakage identification using transient pressure waves provides fast response and enables online monitoring with reduced sensitivity to external conditions [32]. However, routine operational actions such as valve regulation and pump start/stop can also induce transient disturbances, yielding partially overlapping signatures with leakage-induced pressure-wave responses in the measured signals. This similarity poses stringent requirements on the robustness of identification methods.

To address this challenge, we developed an integrated DHN experimental testbed that enables controllable indoor experiments under multi-source, multi-loop, and multi-condition scenarios. The testbed was installed in an indoor space of **4.5 m × 5 m × 3 m** and configured as a three-dimensional, multi-source ring-type pipeline network (**Error! Reference source not found.**). It comprises a heat-source unit, an experimental pipeline network, and an online data-acquisition unit. Two electric boilers operate independently or jointly, while a controllable pump on the return line supports hydraulic balancing and emulates operational variations. Representative hydraulic and thermal conditions (single/multiple heat sources, single/multiple loops, and ring/branched configurations) were reproduced by adjusting heat-source combinations and loop-valve states. Pressure signals were sampled at 2000 Hz for 10 s per trial and used for subsequent preprocessing, feature extraction, and model training and evaluation.

 doi:10.20944/preprints202603.1448.v1

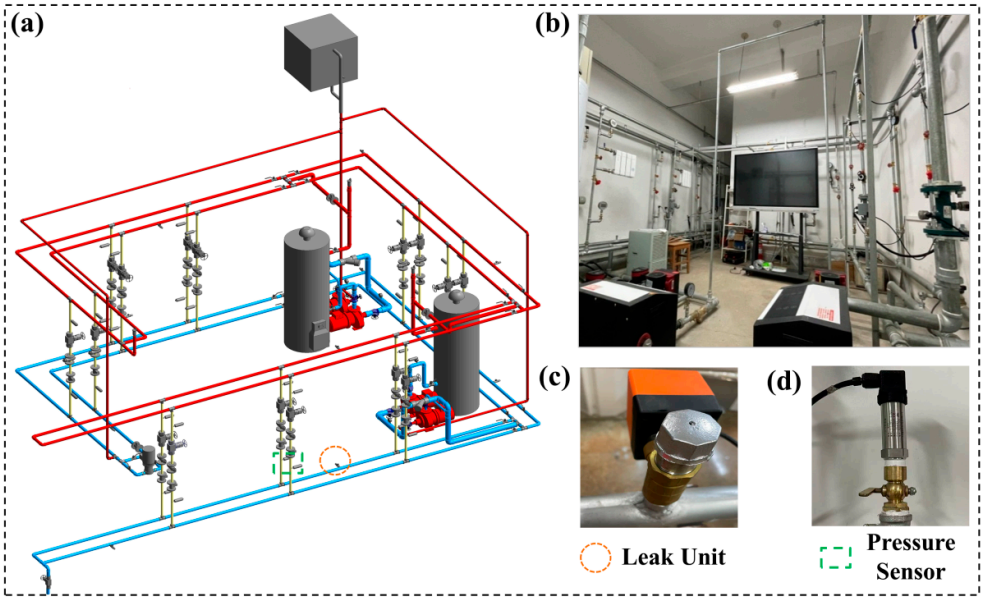


Figure 1. Experimental platform for integrated heating network: (a) Three-dimensional schematic of the pipeline network; (b) Physical diagram of the experimental platform; (c) Leak unit; (d) Pressure sensor

2.2. Dataset Construction and Preprocessing

The constructed pressure-wave dataset covers four representative scenarios in district heating networks, including both leakage events and operation-related disturbances: leakage, valve regulation, pump speed regulation, and emergency valve shut-off. For each scenario, 50 samples were collected, yielding a balanced dataset of 200 samples. All samples were acquired under controllable indoor conditions. By adjusting heat-source combinations, loop-valve states, and return-pump operating modes, multi-source, multi-loop, and multi-condition scenarios were systematically reproduced so that the dataset captures both event-related signatures and the disturbance effects associated with varying operating conditions.

To improve robustness and ensure comparability across trials, the raw pressure records were preprocessed prior to feature extraction. First, a Hampel filter was applied to remove isolated impulsive spikes, reducing the influence of occasional pulse-like artifacts on transient features. Next, to determine an appropriate passband, one-sided power spectral densities (PSDs) were estimated for all records using the Welch method, and the median PSD across trials was used to represent the typical spectrum. A high-frequency noise baseline was then estimated to construct the total cumulative energy curve and the excess-energy cumulative curve (Error! Reference source not found.). The 0.5–300 Hz band contains approximately 77.9% of the total spectral energy while accounting for about 95.1% of the excess energy above the noise baseline, indicating that the primary informative components are concentrated in this low-frequency range. Accordingly, a fourth-order Butterworth band-pass filter with a passband of 0.5–300 Hz was applied, and zero-phase filtering was implemented via forward–backward filtering (filtfilt) to avoid phase distortion that could affect subsequent feature extraction and identification. The resulting pressure-wave signals were used as the unified inputs for subsequent feature construction and model development.

 doi:10.20944/preprints202603.1448.v1

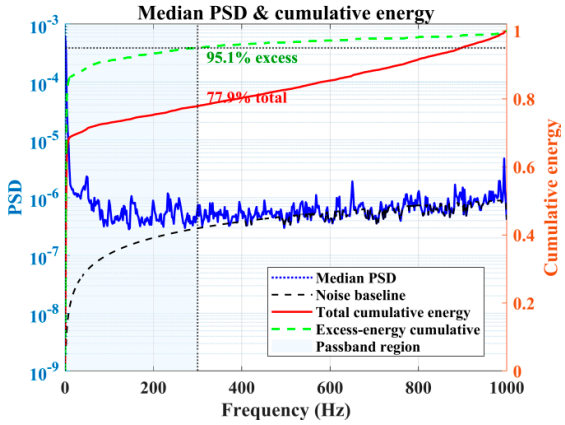


Figure 2. Median PSD and cumulative energy curves for passband selection.

3. Methodology

As summarized in Figure 3, the proposed framework starts from multi-condition transient pressure-wave acquisition in an integrated district heating network testbed. The raw pressure signals are first preprocessed to suppress outliers and noise, followed by extracting handcrafted statistical descriptors in both the time and frequency domains to form candidate feature vectors. A multiclass ReliefF scheme is then employed to rank features by their discriminative relevance and to select a compact subset for downstream learning. To mitigate the small-sample limitation under multi-condition settings, a class-wise GAN is trained in the selected feature space to generate synthetic feature vectors, which are used to augment the training set only. Finally, multiple classifiers are trained on the augmented feature set and used to identify operating conditions and leakage-related events on real test samples.

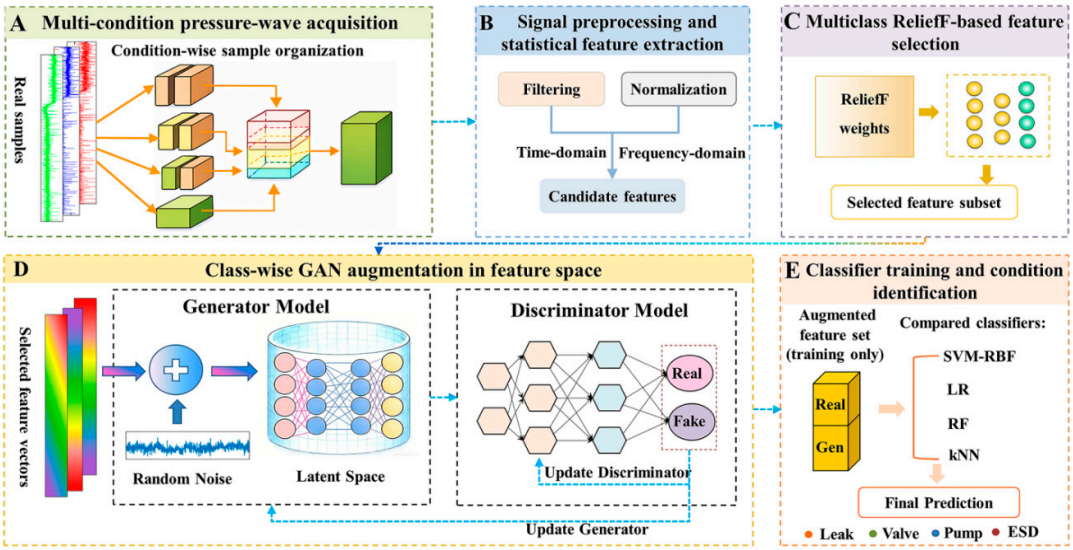


Figure 3. Overall workflow of the proposed DHN transient pressure-wave based condition identification framework.

3.1. Feature Extraction

To represent transient pressure wave signals under multi-condition operations in a compact and interpretable form, we extract a set of hand-crafted features from each preprocessed segment. Time-

 doi:10.20944/preprints202603.1448.v1

domain statistical and shape-related descriptors have been widely adopted in pipeline leak diagnosis to capture amplitude level, dispersion, and impulsiveness, while frequency-domain descriptors summarize spectral energy distribution and characteristic frequencies that are sensitive to leakage-related acoustic or pressure-wave signatures [33–35]. For district heating applications, the leakage signature can vary with operating conditions such as temperature and pressure, which further motivates a robust feature representation that remains discriminative under multi-condition transients [36,37]. Accordingly, we construct a candidate feature pool by combining time-domain and frequency-domain descriptors. The complete definitions of the time-domain features (P1–P14) and the frequency-domain features (P15–P21) are provided in **Table 1** and T**able 2,** respectively.

Table 1. Time-domain features**.**

ID	Feature Name	Formula	ID	Feature Name	Formula
P1	Mean	$\bar{p} = \frac{1}{N}\sum_{i=1}^{N} p_i$	P8	Peak-to-peak value	$p_{p-p} = p_{\max} - p_{\min}$
P2	Absolute mean	$\overline{p_{abs}} = \frac{1}{N}\sum_{i=1}^{N} \lvert p_i \rvert$	P9	Skewness coefficient	$\alpha = \frac{\frac{1}{N}\sum_{i=1}^{N}(p_i - \bar{p})^3}{\sigma^3}$
P3	Root mean square	$p_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} p_i^2}$	P10	Kurtosis coefficient	$\beta = \frac{\frac{1}{N}\sum_{i=1}^{N}(p_i - \bar{p})^4}{\sigma^4}$
P4	Square root magnitude	$p_{srm} = \left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert p_i \rvert}\right)^2$	P11	Crest factor	$C = \frac{p_p}{p_{rms}}$
P5	Zero Crossing	$Z_C = \frac{1}{2}\sum_{n=2}^{N}\lvert \mathrm{sgn}(p_i) - \mathrm{sgn}(p_{i-1}) \rvert$	P12	Impulse factor	$I = \frac{p_p}{\bar{p}}$
P6	Standard deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - \bar{p})^2}$	P13	Margin Factor	$L = \frac{p_p}{p_{smr}}$
P7	Peak value	$p_p = \max \lvert p_i \rvert$	P14	Shape factor	$S = \frac{p_{rms}}{\overline{p_{abs}}}$

Note: sgn(·) is the sign function, defined as sgn(x)=1 if x ≥ 0 and sgn(x) = −1 otherwise.

Table 2. Frequency-domain features**.**

ID	Feature Name	Formula	ID	Feature Name	Formula
P15	Mean spectral power	$x_{\mathrm{MSP}} = \frac{1}{N}\sum_{i=1}^{N} w(f_i)$	P19	Frequency variance	$x_{\mathrm{FV}} = \frac{\sum_{i=1}^{N}(f_i - x_{\mathrm{FC}})^2 w(f_i)}{\sum_{i=1}^{N} w(f_i)}$
P16	Spectral centroid	$x_{\mathrm{SC}} = \frac{\sum_{i=1}^{N} f_i w(f_i)}{\sum_{i=1}^{N} w(f_i)}$	P20	Spectral entropy	$H_{\mathrm{SE}} = -\sum_{i=1}^{N_f} \tilde{w}(f_i)\ln \tilde{w}(f_i)$
P17	Mean square frequency	$x_{\mathrm{MSF}} = \frac{\sum_{i=1}^{N} f_i^2 w(f_i)}{\sum_{i=1}^{N} w(f_i)}$	P21	Spectral flatness	$x_{\mathrm{SF}} = \frac{\exp\left(\frac{1}{N_f}\sum \ln\left(w(f_i)+\epsilon\right)\right)}{\frac{1}{N_f}\sum w(f_i)}$
P18	Peak frequency	$x_{\mathrm{PF}} = \arg\max_{f_i} w(f_i)$			

 doi:10.20944/preprints202603.1448.v1

Note: fi (i=1,…,Nf) denotes discrete frequency bins on the one-sided spectrum [0, fs/2], where Nf is the number of bins and fs is the sampling rate. w(fi) is the spectral weight at fi, defined as the one-sided power spectral density (PSD) of the signal. The normalized weight is $\tilde{w}(f_i) = w(f_i) / \sum_{j=1}^{N_f} w(f_j)$, which is used in spectral entropy. For spectral flatness (SFM), a small constant ε (e.g., 10−12) is added to avoid numerical issues in ln(·).

3.2. Multiclass ReliefF-based Feature Selection

To construct a compact and interpretable discriminative input from multi-condition pressure-wave features, we adopt the multiclass extension of ReliefF to evaluate and rank candidate features. ReliefF is a filter-based feature selection method that updates feature weights by comparing the differences between near-hit (same-class) neighbors and near-miss (different-class) neighbors within a local neighborhood. It is decoupled from the downstream classifier, computationally efficient, and sensitive to local discriminative information, making it suitable for rapid screening and dimensionality reduction of engineering statistical features [27,29,38].

Let the training set contain N samples, each described by P candidate features. The feature vector of the i-th sample is denoted by $p_i \in \mathbb{R}^P$, with class label class(p_i)$\in \{1,\cdots,C\}$. ReliefF initializes the feature-weight vector as $\boldsymbol{W}^{(0)} = (W_1^{(0)},\cdots,W_P^{(0)})^\top = 0$, and performs T random iterations (here T = 500). At iteration t, a reference sample p_i is randomly selected. Within the same class, the k nearest neighbors of p_i (here k = 10) are denoted as $H_j(p_i)$ (j=1,…,k). Meanwhile, for each different class c≠class(p_i), the k nearest neighbors of p_i within class c are denoted as $M_{c,j}(p_i)$ (j=1,…,k). Feature weights are then updated according to the principle of "penalizing within-class differences and rewarding between-class differences", with class-prior ratios incorporated for multiclass aggregation. For the r-th feature (r=1,…,P), let A_r denote the corresponding feature mapping (i.e., $A_r(p) = p_r$). The weight update is given by

$$W_r^{(t+1)} = W_r^{(t)} - \frac{1}{Tk}\sum_{j=1}^{k}\mathrm{diff}(A_r, p_i, H_j(p_i)) + \sum_{c\neq \mathrm{class}(p_i)}\left[\frac{P_c}{1-P_{\mathrm{class}(p_i)}}\cdot\frac{1}{Tk}\sum_{j=1}^{k}\mathrm{diff}(A_r, p_i, M_{c,j}(p_i))\right] \tag{1}$$

where $W^{(t)}$ is the weight vector after th t-th iteration and $W^{(t+1)}$ is the updated weight vector. P_c denotes the prior proportion of class c in the training set (estimated by class frequency), and $P_{\mathrm{class}(p_i)}$ is the prior proportion of the class to which the reference sample belongs. This update mechanism assigns higher weights to features that are more stable within the same-class neighborhood and more discriminative across different classes, thereby reflecting their local discriminative contribution under multi-condition identification.

For continuous features, a normalized difference metric is adopted:

$$\mathrm{diff}(A, p_i, p_\ell) = \begin{cases} \dfrac{|A(p_i) - A(p_\ell)|}{\max(A) - \min(A)}, A \text{ is the continuous feature} \\ \mathbb{I}[A(p_i) \neq A(p_\ell)], A \text{ is the discrete feature} \end{cases} \tag{2}$$

where p_i and p_ℓ denote any two samples, $\mathbb{I}[\bullet]$ is the indicator function. To avoid information leakage, max(A) and min (A) are computed using the training set only and are kept fixed during validation/testing. After T iterations, the final weight vector $W^{(T)}$ is obtained. Features are ranked in descending order of $W_r^{(T)}$, and the Top-q features are retained to form the final feature subset for subsequent classifier training and evaluation. The value of q is determined within the training set (e.g., based on cross-validation performance or an elbow criterion on the weight curve), and the test set is used only for final generalization evaluation.

 doi:10.20944/preprints202603.1448.v1

3.3. Data Augmentation Using GAN

Due to the limited number of controllable experiments, the multi-condition pressure-wave dataset is relatively small, which may restrict the ability of supervised models to learn stable decision boundaries under operation-related transients. To mitigate this issue, we employ generative adversarial networks (GANs) for data augmentation in the feature space. GAN-based augmentation has been increasingly used for small-sample or imbalanced pipeline leakage and fault diagnosis, demonstrating improved robustness and generalization [20–22].

Given the four-class identification task, we adopt a class-wise augmentation strategy to ensure label consistency. A separate GAN is trained for each class using the corresponding real feature vectors p_i. After training, synthetic feature vectors are generated by sampling latent variables $z_i \sim p_z(z)$ and mapping them through the generator $G(\cdot)$ to obtain $\hat{p}_i = G(z_i)$. In this study, the original dataset contains 200 real samples (50 per class). The dataset size is expanded by a factor of five, increasing the total number of samples from 200 to 1000. To keep the dataset balanced, each class is expanded from 50 to 250 samples, i.e., 200 synthetic samples are generated per class. Importantly, synthetic samples are used only for training-time augmentation, while all evaluations are conducted on real samples to avoid information leakage.

The discriminator D is optimized to assign high confidence to real samples and low confidence to generated samples:

$$\max_D V(D,G) = E_{p_i \sim p_{\text{data}}(p)}[\log D(p_i)] + E_{z_i \sim p_z(z)}[\log(1 - D(G(z_i)))] \tag{3}$$

where $E[\cdot]$ denotes the expectation operator, $p_i \sim p_{\text{data}}(p)$ denotes a real feature vector sampled from the data distribution, $z_i \sim p_z(z)$ denotes a latent variable sampled from the prior distribution, $G(\cdot)$ maps z_i to a synthetic feature vector $G(z_i)$, and $G(\cdot)$ outputs the probability that an input feature vector is real.

With D fixed, the generator G is trained to fool the discriminator by minimizing:

$$\min_G V(D,G) = E_{z_i \sim p_z(z)}[\log(1 - D(G(z_i)))] \tag{4}$$

This alternating optimization between D and G is repeated until convergence or until a predefined number of training epochs is reached.

3.4. Training Protocol and Evaluation Metrics

To ensure a fair evaluation of the proposed framework, all experiments were conducted under a strict training/testing separation protocol. The dataset was first divided into training and test sets, and the test set was used exclusively for final performance evaluation. All data-dependent procedures, including feature normalization, feature ranking, and feature subset selection, were performed using the training set only, and the obtained parameters were then applied to the test set. This protocol prevents information leakage and ensures that the reported results reflect the true generalization capability on unseen real samples.

After signal preprocessing and statistical feature extraction, feature selection was performed on the training set. In addition to the proposed multiclass ReliefF method [27–29], two representative feature selection methods, namely mRMR [24] and LASSO [25], were introduced for comparison. ReliefF evaluates feature importance according to local discriminative information in the neighborhood space and is particularly suitable for multi-class problems under limited-sample conditions [39]. mRMR aims to obtain a compact feature subset by balancing feature relevance and redundancy [24], whereas LASSO performs embedded feature selection through sparse regularization [40]. To evaluate the influence of the downstream learning model, four representative classifiers were further considered, including support vector machine with radial basis function kernel (SVM-RBF) [41,42], logistic regression (LR) [43], random forest (RF) [28], and k-nearest neighbors (kNN) [44,45]. These models were selected because they represent kernel-based, linear, ensemble-based, and distance-based classification paradigms, respectively, and have been widely

 doi:10.20944/preprints202603.1448.v1

used in data-driven fault diagnosis and pattern recognition studies. Detailed mathematical derivations of these methods are beyond the scope of this paper, and interested readers may refer to the corresponding literature for further details.

Because the original dataset is limited in size, GAN-based augmentation in feature space was introduced to enrich the training distribution. Augmentation was applied only to the training set, whereas the test set always consisted of real measured samples. Therefore, any performance gain observed after augmentation can be interpreted as an improvement in model robustness and generalization rather than an artifact of synthetic-sample reuse. The classifier performance before and after augmentation is denoted as Base and Augmented, respectively. To reduce the influence of randomness in feature selection, classifier training, and synthetic sample generation, each experiment was repeated multiple times under different random seeds, and the averaged results were reported. All comparative experiments were conducted under the same protocol to ensure consistency and fairness.

For the multi-condition identification task, the classification performance was evaluated using metrics derived from the confusion matrix [46,47]. For class i, TP_i , FP_i, FN_i, and TN_i denote the numbers of true positives, false positives, false negatives, and true negatives, respectively. The overall classification accuracy is defined as

$$\text{Accuracy} = \frac{\sum_{i=1}^{C} TP_i}{N_{\text{ts}}} \tag{5}$$

where C is the number of classes and N_{ts} is the total number of test samples.

To assess the balanced performance across all operating conditions, the macro-averaged F1-score was adopted:

$$\text{Macro-F1} = \frac{1}{C}\sum_{i=1}^{C} \text{F1}_i \tag{6}$$

The class-wise F1-score is calculated as

$$\text{F1}_i = \frac{2 \cdot \text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{7}$$

with

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i}, \text{Recall}_i = \frac{TP_i}{TP_i + FN_i} \tag{8}$$

Since leakage recognition is the primary engineering concern in this study, the performance of the leakage class was further evaluated separately. For the leakage class, TPleak, FPleak , FNleak , and TNleak are defined analogously. The leakage-class precision, recall, specificity, and F1-score, denoted as **Leak-Prec**, **Leak-Rec**, **Leak-Spec, and Leak-F1**, respectively, are given by

$$\text{Leak-Prec} = \frac{TP_{\text{leak}}}{TP_{\text{leak}} + FP_{\text{leak}}} \tag{9}$$

$$\text{Leak-Rec} = \frac{TP_{\text{leak}}}{TP_{\text{leak}} + FN_{\text{leak}}} \tag{10}$$

$$\text{Leak-Spec} = \frac{TN_{\text{leak}}}{TN_{\text{leak}} + FP_{\text{leak}}} \tag{11}$$

$$\text{Leak-F1} = \frac{2 \cdot \text{LeakPrec} \cdot \text{LeakRec}}{\text{LeakPrec} + \text{LeakRec}} \tag{12}$$

 doi:10.20944/preprints202603.1448.v1

In this study, **Accuracy** and **Macro-F1** were used to evaluate the overall identification performance across all operating conditions, whereas **Leak-Prec**, **Leak-Rec**, **Leak-Spec**, and **Leak-F1** were used to specifically quantify the distinguishability of leakage events from operation-induced disturbances. For each classifier and feature selection method, the results under both Base and Augmented settings were compared on the same real test set. This evaluation protocol provides a comprehensive assessment of classifier effectiveness, feature selection rationality, and the contribution of GAN-based feature-space augmentation.

4. Results

4.1. Feature Selection Results

To obtain input features with stronger discriminative capability for multi-condition pressure-wave identification, we employ multiclass ReliefF to evaluate the importance of candidate features. To avoid information leakage, both the ReliefF weight estimation and feature selection are performed exclusively on the training set. ReliefF exploits the local neighborhood structure of samples by quantifying feature-wise differences between near-hit (same-class) and near-miss (different-class) neighbors and updating feature weights accordingly; a larger weight indicates a greater contribution to class separability. Because the random sampling in ReliefF can introduce variability in the estimated weights and the number of neighbors k affects the stability of neighborhood estimation, we conduct a sensitivity analysis over different k values. For each k, we repeat the procedure 20 times using different random seeds and report the mean weight (with the standard deviation) as the final estimate of feature importance. The results indicate that the weight curves and rankings become stable when k = 15; therefore, k = 15 is adopted in the subsequent experiments. The final feature weights are shown in Figure 4, based on which a subset of high-weight features is selected for downstream classification

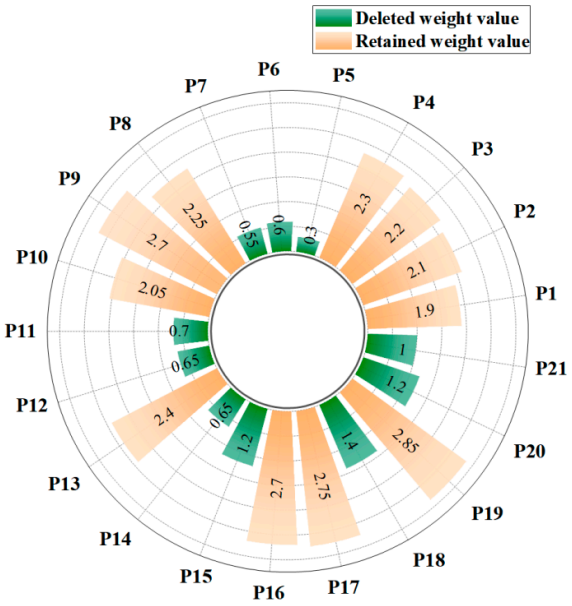


Figure 4. Weight values for all features.

Feature weights quantify the discriminative contribution of each candidate feature to multiclass identification: a larger weight indicates that the feature better separates samples from different classes within local neighborhoods and thus contributes more to the classification decision. As shown in Figure 4, ReliefF assigns importance scores to all candidate features (P1–P21), where the orange sectors denote retained features and the green sectors denote discarded ones. The results indicate that

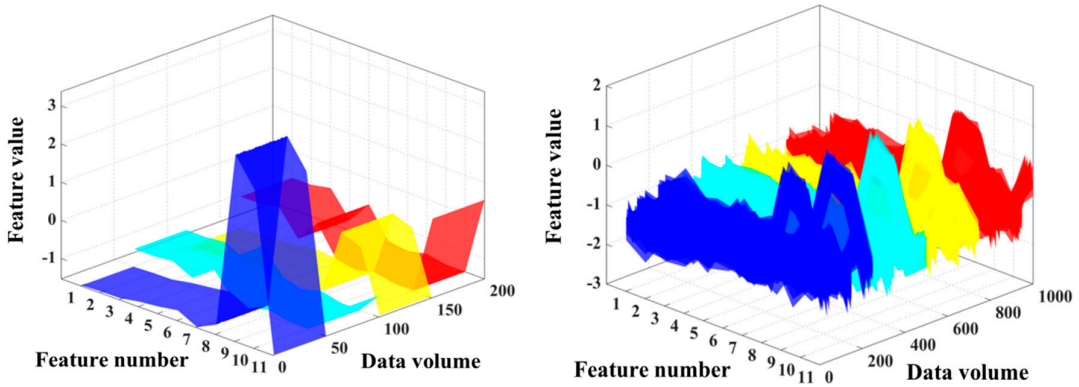
 doi:10.20944/preprints202603.1448.v1

P5–P7, P11, P12, P14, P15, P18, P20, and P21 receive consistently low weights, suggesting limited discriminative power (likely due to redundancy with higher-weight features or insufficient sensitivity to leakage transients under operational disturbances); these features are therefore removed. Consequently, the 11 highest-ranked features (P1–P4, P8–P10, P13, P16, P17, and P19) are retained to form the final feature vector for each signal segment. This compact subset jointly covers key time-domain statistical and frequency-domain spectral characteristics, preserving the principal discriminative information while reducing input dimensionality, thereby providing a robust and interpretable representation for subsequent data augmentation and classifier training.

4.2. Quality Assessment of GAN-Generated Samples

The pressure-wave dataset includes four operating conditions: leakage, valve regulation, pump speed regulation, and emergency valve shut-off. Each class contains 50 samples, yielding 200 samples in total. For each sample, 11 statistical features are extracted, resulting in an original feature matrix $\boldsymbol{P} \in \mathbb{R}^{200\times11}$. To alleviate the small-sample issue, a generative model is used to synthesize an additional $5N$ samples during training (N = 200) to augment the training data. To ensure consistent feature scaling, all features are standardized using Z-score normalization before being fed into the generative model/discriminator.

Figure 5 compares the feature-space distributions of the real and GAN-generated samples. Figure 5a shows the standardized distributions of the 200 original samples over the 11 extracted features, whereas Figure 5b displays the corresponding distributions of 1000 synthetic samples generated in the same feature space. For the leakage class, the generated samples exhibit stronger fluctuations in features P9 and P11 (high-frequency-related features), and they also show a wider range of variation in amplitude-related features (e.g., RMS and peak-to-peak value). Since leakage signals typically involve higher amplitude fluctuations and pronounced impulsiveness, these characteristics are amplified in the generated data. Compared with valve regulation and pump speed regulation, leakage samples show more pronounced variations in high-frequency-related features, and the generated data further expands this variation range, which can help the model distinguish leakage more effectively. In addition, the impulsiveness of leakage signals in energy-related features (e.g., margin factor and kurtosis coefficient) is also enhanced, making these features more salient for identifying leakage.



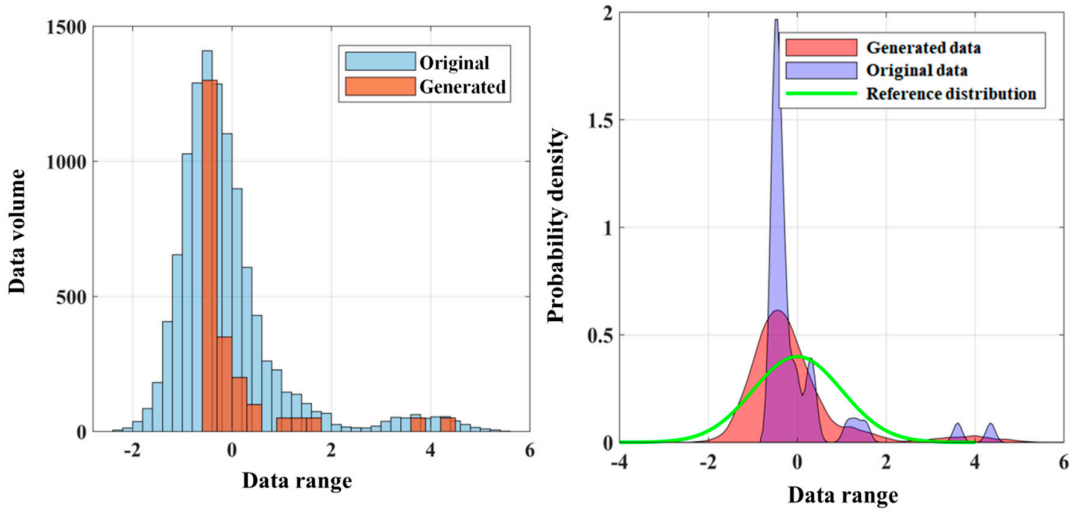
(**a**) Original data (**b**) Generated data

Figure 5. Three-dimensional distribution of data features.

Overall, the original data occupy a relatively concentrated region in the feature space, whereas the generated data exhibit broader coverage and stronger diversity. By retaining the key characteristics of the original data and increasing the dispersion of leakage samples in extreme feature regions, the augmented set provides more informative examples for learning decision boundaries and improves leakage separability.

 doi:10.20944/preprints202603.1448.v1

Figure 6 compares the value-range distributions of the GAN-generated data and the original data. As shown in Figure 6a, the generated data follow the original distribution closely within the main interval (e.g., from −2 to 2). Minor discrepancies are observed in the tail regions, where the generated samples provide broader coverage due to the increased sample size, effectively enriching the feature distribution. The probability density comparison in Figure 6b further contrasts the original data, the generated data, and a reference distribution (green curve). The generated data align well with the reference distribution around zero, indicating that the core statistical characteristics of the original data are reproduced effectively. Although deviations remain in the extreme tails, the generated data cover the edge regions more extensively, complementing portions that are sparsely represented in the original dataset. This suggests that the generated data preserve similarity to the original samples while enhancing coverage of extreme cases, which can improve model robustness to rare conditions. In particular, leakage-related descriptors, especially magnitude-related and spectral-related features, typically show larger dispersion and heavier tails than transients caused by routine control actions, which enhances separability in the feature space.

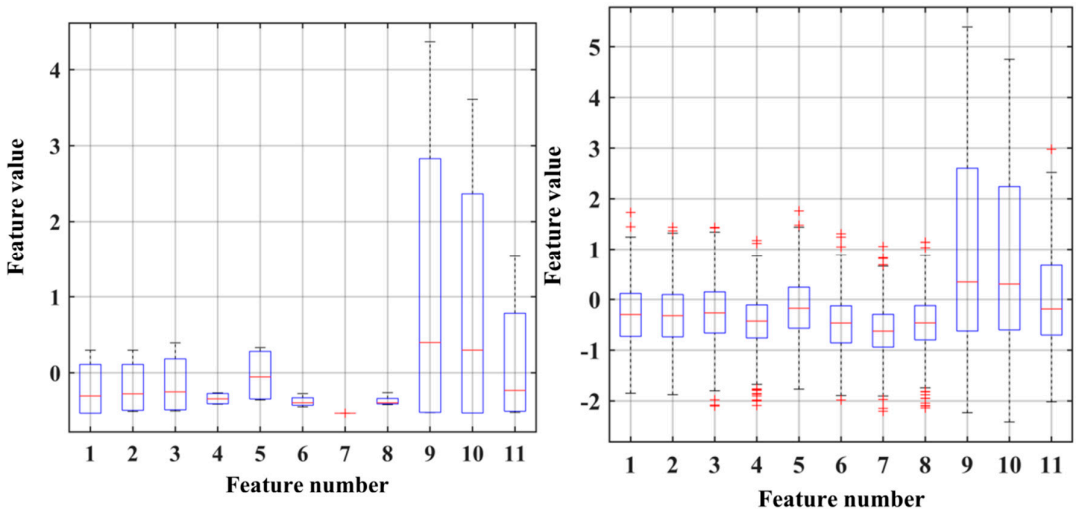


(**a**) Histogram comparison (**b**) Probability density comparison

Figure 6. Distribution comparison of original and generated data**.**

Figure 7 shows the boxplot distributions of the original data and the generated data across the 11 features. Each boxplot displays the interquartile range, maximum and minimum values, and outliers (if any) for a given feature. In Figure 7a, the original data exhibit relatively large fluctuations in some features (e.g., P9 and P10), with feature values spanning roughly from 0 to 4 and including a small number of outliers. Figure 7b presents the distributions of the generated features. Compared with the original data, the generated data not only reproduce the main distribution patterns of the features, but also provide a denser and more complete coverage of the feature space by increasing the sample size. Notably, larger variations are observed in features P8 and P11, indicating that the generated samples capture broader fluctuations in leakage-related characteristics. Overall, the generated data enrich the training set by expanding the feature distribution support, which helps improve the model's generalization to extreme cases. Leakage signals tend to exhibit wider ranges in the boxplots, particularly for peak-to-peak value, margin factor, and root mean square, showing more prominent dispersion and occasional extreme values, whereas valve-regulation and pump-speed-regulation signals are generally more stable with smaller variability.

 doi:10.20944/preprints202603.1448.v1



(**a**) Original data (**b**) Generated data

Figure 7. Boxplots of original and generated data.

Figure 8**1** presents probability–probability (P–P) plots that compare the probability14 distributions of the generated data and the original data across different features. The plots are constructed using normalized feature values; the horizontal axis represents the cumulative probability (from 0 to 1), and the vertical axis represents the corresponding quantiles. Eleven curves are shown, each corresponding to one of the 11 features. When a curve lies close to the diagonal line, it indicates that the generated and original data follow similar statistical distributions for that feature.

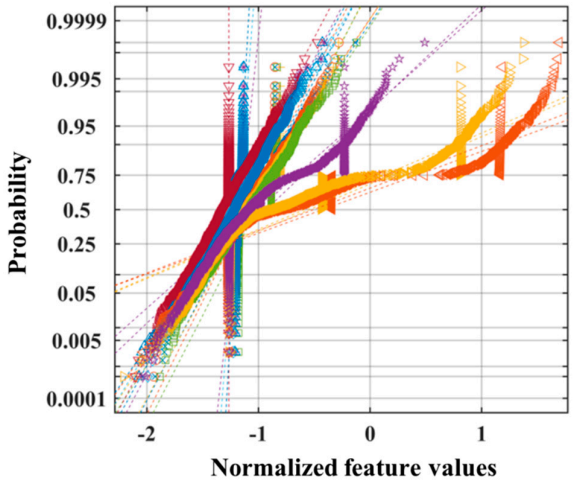


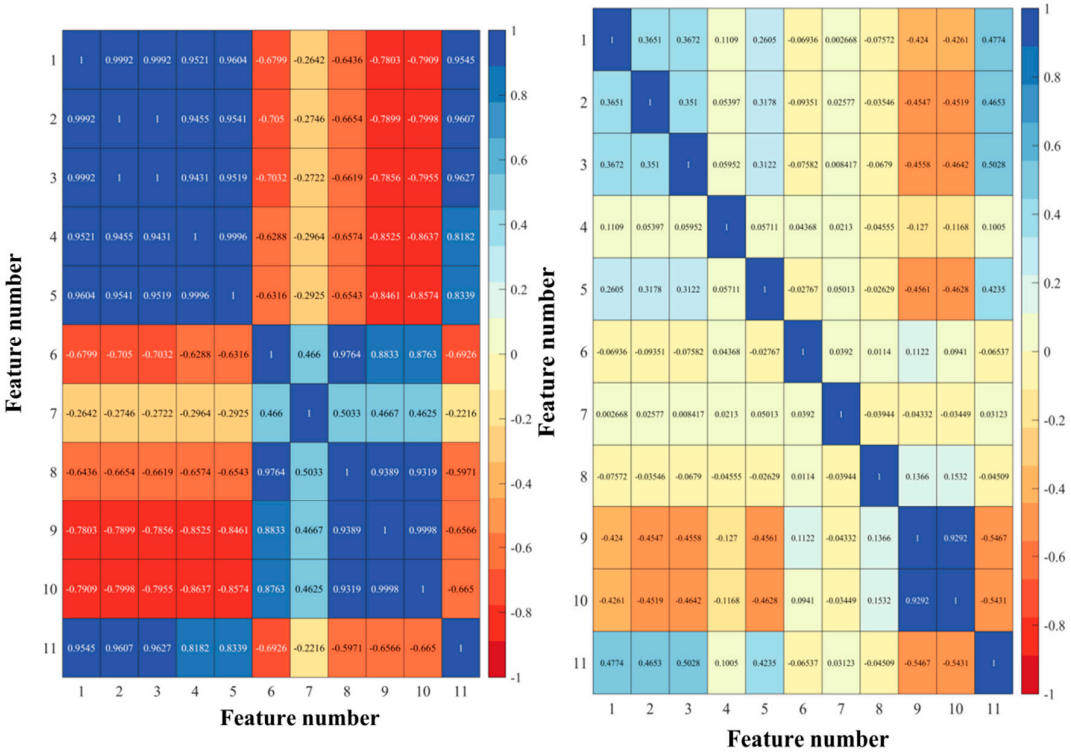
Figure 8. Probability–probability (P–P) plots of generated and original data.

Overall, most features exhibit good agreement in the central region (e.g., from approximately −1 to 1), suggesting that the generated data closely match the original data in the main body of the distribution. However, in the extreme regions, noticeable deviations appear for some features, indicating differences in tail behavior and that the generated data may not fully replicate the boundary characteristics of the original data. In particular, leakage-related features show high similarity between the generated and original distributions within the core region, implying that the generated samples preserve the main characteristics of leakage signals. Nevertheless, for leakage-related frequency features, discrepancies in the extreme tails may still exist compared with valve-regulation or pump-speed-regulation signals, especially in the tail segments of the distributions.

 doi:10.20944/preprints202603.1448.v1

Figure 9 shows the feature correlation matrices of the original data and the generated data, which are used to analyze correlations among the 11 extracted features. The correlation coefficient ranges from −1 (perfect negative correlation) to 1 (perfect positive correlation). In the heatmaps, darker colors indicate stronger correlations; red denotes positive correlation, blue denotes negative correlation, and lighter colors indicate weak or no correlation. In Figure 9a, the original data exhibit a strong positive correlation (close to 1) between Feature 9 (spectral centroid) and Feature 10 (mean square frequency), because both features describe the frequency distribution of the signal and characterize the tendency of spectral energy concentration in the frequency domain. Features 6 (kurtosis coefficient) and 7 (skewness coefficient) show a pronounced negative correlation (approximately −0.7 to −0.9) with amplitude-related features such as Feature 3 (root mean square) and Feature 5 (square root amplitude), indicating that as impulsiveness (kurtosis coefficient) increases, the overall signal energy (root mean square) and amplitude-related measures tend to decrease. This may reflect an inherent trade-off between concentrated impulsive fluctuations and the overall energy level. In Figure 9b, the generated data largely preserve a correlation structure similar to that of the original data, especially the relationships between frequency-related features and energy-related features. However, the correlations among some feature pairs become weaker. In particular, the negative correlation between Feature 1 (mean) and frequency-related features is reduced, suggesting that the coupling between leakage-related impulsiveness and spectral characteristics may be less pronounced in the generated samples. Moreover, for leakage signals, the generated data tend to show stronger fluctuations in amplitude-related or energy-related features, making indicators such as kurtosis coefficient, skewness coefficient, and margin factor more salient. In contrast, valve-regulation and pump-speed-regulation signals exhibit noticeably different and weaker inter-feature correlations, reflecting their distinct transient behaviors from leakage signals.

Although the above analyses confirm the distributional plausibility of the GAN-generated feature vectors under the adopted fivefold augmentation setting, the final identification performance may still depend on the choice of classifier, augmentation ratio, and feature-input scheme. Therefore, these factors are further examined in Section 4.3 through a series of comparative experiments.



(**a**) Original data (**b**) Generated data

 doi:10.20944/preprints202603.1448.v1

Figure 9. Feature correlation matrices of original and generated data.

4.3. Classification Performance and Comparison Studies

After verifying the quality of the GAN-generated feature vectors, further experiments were carried out to evaluate the final multi-condition identification performance. In this section, three aspects are compared systematically: the effectiveness of different classifiers, the influence of augmentation ratio, and the impact of different feature-input schemes. The objective is not only to identify the most suitable model configuration for the present dataset, but also to clarify the respective roles of classifier design, synthetic sample size, and feature selection in the proposed framework.

4.3.1. Benchmarking of Candidate Classifiers

To determine a suitable classifier for the subsequent comparison studies, four representative supervised learning models were benchmarked, including RF, SVM-RBF, LR, and kNN. These models were selected because they represent distinct learning mechanisms, namely ensemble-based tree learning, kernel-based classification, linear regularized modeling, and distance-based classification. For consistency, all classifiers were trained and tested under the same data partition and evaluation protocol, using the ReliefF-selected feature subset as input.

Table 3 lists the optimized hyperparameter settings of the candidate classifiers used in this study. These settings were determined under the same experimental protocol to ensure a fair comparison among different learning models.

Table 3. Optimized hyperparameter settings of the candidate classifiers.

Classifier	Optimized hyperparameter setting
SVM-RBF	C = 10, kernel = rbf, gamma = 0.05556, KernelScale = 3
LR	Regularization = ridge, Lambda = 0.001
RF	Criterion = deviance, n_estimators = 200, MaxNumSplits = 60, MinLeafSize = 10, NumVariablesToSample = 2
kNN	k = 7, distance = euclidean

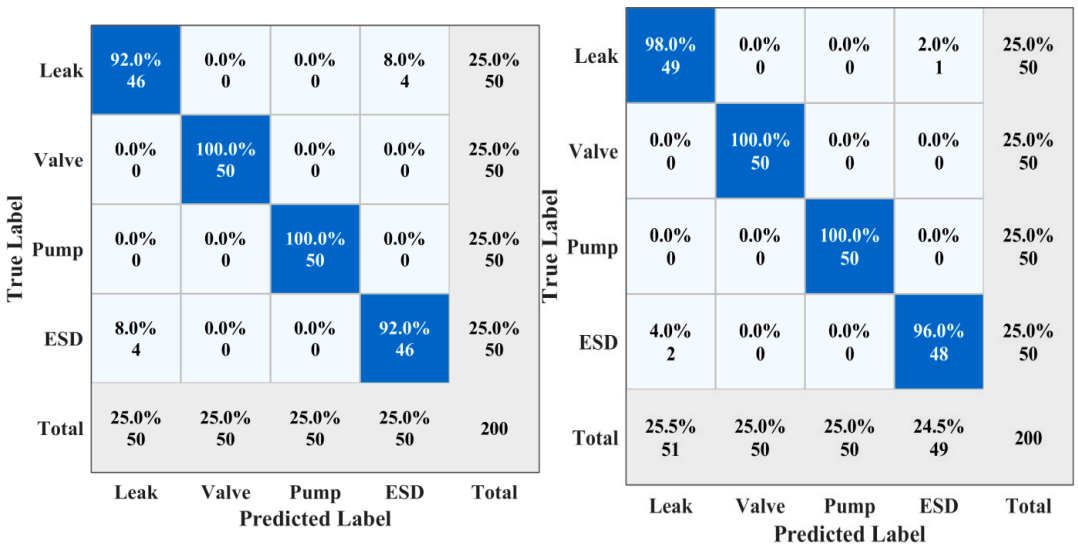
Table 4 reports the comparative performance of four representative classifiers (SVM-RBF, LR, RF, and kNN) evaluated on real test samples under an outer 5-fold cross-validation protocol. To quantify both overall discrimination and leakage-oriented detection capability, we report Accuracy, Macro-F1, and leakage-specific metrics including Leak-F1, Leak-Prec, Leak-Rec, and Leak-Spec. Results are summarized as mean ± standard deviation across the outer folds, ensuring that the reported statistics reflect both average performance and cross-split robustness. Overall, the augmented setting consistently improves the leakage-related metrics—particularly recall and F1—indicating that the generated feature vectors can enrich the training distribution and enhance generalization to noisy real-world measurements.

 doi:10.20944/preprints202603.1448.v1

Table 4. Classification performance of four classifiers with original-data training and GAN-generated data augmentation (training only).

Classifier	Acc (Qrig)	Macro-F1 (Qrig)	Leak-F1 (Qrig)	Leak-Prec (Qrig)	Leak-Rec (Qrig)	Leak-Spec (Qrig)	Acc (Gen)	Macro-F1 (Gen)	Leak-F1 (Gen)	Leak-Prec (Gen)	Leak-Rec (Gen)	Leak-Spec (Gen)
SVM-RBF	0.925 ± 0.047	0.924 ± 0.047	0.849 ± 0.089	0.870 ± 0.131	0.840 ± 0.114	0.973 ± 0.028	0.970 ± 0.033	0.970 ± 0.033	0.942 ± 0.082	0.945 ± 0.050	0.900 ± 0.141	0.980 ± 0.018
LR	0.970 ± 0.027	0.970 ± 0.027	0.942 ± 0.052	0.927 ± 0.076	0.960 ± 0.055	0.973 ± 0.028	0.980 ± 0.011	0.980 ± 0.011	0.961 ± 0.022	0.945 ± 0.050	0.980 ± 0.045	0.980 ± 0.018
RF	0.960 ± 0.014	0.960 ± 0.014	0.920 ± 0.029	0.928 ± 0.072	0.920 ± 0.084	0.973 ± 0.028	0.985 ± 0.014	0.985 ± 0.014	0.970 ± 0.027	0.964 ± 0.050	0.980 ± 0.045	0.987 ± 0.018
kNN	0.925 ± 0.040	0.925 ± 0.040	0.850 ± 0.082	0.845 ± 0.075	0.860 ± 0.114	0.947 ± 0.030	0.940 ± 0.045	0.939 ± 0.047	0.873 ± 0.106	0.905 ± 0.098	0.860 ± 0.167	0.967 ± 0.041

Figure 10 compares the class-wise prediction patterns of the RF classifier via confusion matrices aggregated on real test samples. When trained on original data only (Figure 10a), leakage events are occasionally confused with operational transients, reflecting the similarity between leakage-induced pressure-wave signatures and routine disturbances. After augmenting the training set with GAN-generated synthetic feature vectors (Figure 10b), the confusion matrix exhibits stronger diagonal dominance and fewer leakage-related misclassifications, indicating improved separability and robustness under noisy measurements.



(**a**) Original data (**b**) Generated data

Figure 10. Confusion-matrix analysis of the RF classifier under original-data training and GAN-generated data augmentation (training only).

 doi:10.20944/preprints202603.1448.v1

4.3.2. Sensitivity Analysis of Augmentation Ratio

After selecting the representative classifier, we further conducted a sensitivity analysis with respect to the augmentation level. Although a fivefold expansion was used in Section 4.2 to assess the quality of the generated samples, this ratio is not necessarily optimal for downstream classification. We therefore varied the augmentation ratio *r* (i.e., the target number of training samples per class after augmentation relative to the original training size) and evaluated performance on real test samples under outer 5-fold cross-validation. This analysis clarifies how the amount of GAN-generated synthetic feature vectors affects generalization, thereby separating mere distributional plausibility of generated data from their practical utility in improving leakage identification.

Figure 11 indicates that the augmentation ratio r exerts a **non-monotonic** influence on generalization performance. As r increases from 1 to moderate levels, both Macro-F1 and Leak-Rec generally improve, suggesting that GAN-generated feature vectors can enrich the training distribution and strengthen leakage-related separability under noisy measurements. When r becomes overly large, the gains saturate and may slightly degrade, implying that excessive synthetic data can introduce distributional bias and reduce the effective contribution of real samples. These results highlight the necessity of selecting a suitable augmentation level rather than assuming that more synthetic data always yields better generalization.

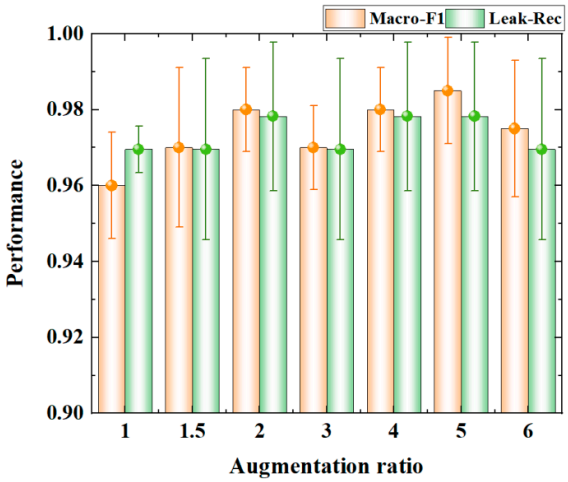


Figure 11. Sensitivity of classification performance to the augmentation ratio of GAN-generated synthetic feature vectors.

4.3.3. Validation of the Feature-Selection Strategy

To further verify the rationality of adopting ReliefF in the proposed framework, additional comparisons were conducted among different feature-input schemes. Besides the ReliefF-selected compact subset, the classifier was also evaluated using the full feature set and other representative feature-selection strategies. Table 5 compares the proposed ReliefF-based feature-selection strategy with two representative alternatives, namely LASSO and mRMR, in terms of both the **selected feature subsets** and the resulting **classification performance**. All methods are evaluated under the same outer 5-fold cross-validation protocol on real test samples, and metrics are summarized as mean ± standard deviation. Besides overall accuracy and Macro-F1, leakage-oriented metrics (Leak-F1, Leak-Prec, Leak-Rec, and Leak-Spec) are reported to explicitly quantify the capability of detecting leakage events while suppressing false alarms. As shown in **Table 5**, the ReliefF-selected subset achieves the best overall performance and consistently higher leakage-related scores, indicating that the selected features are more discriminative for separating leakage-induced pressure-wave signatures from operational disturbances. This comparison supports the rationality of adopting ReliefF to derive a compact yet informative feature subset in the proposed framework

 doi:10.20944/preprints202603.1448.v1

Table 5. Performance comparison of different feature-selection strategies and the resulting selected feature subsets.

Classifier	Selected features	Acc	Macro -F1	Leak -F1	Leak -Prec	Leak -Rec	Leak -Spec
LASSO	P1, P3, P5, P11, P13–P16, P18	0.980 ± 0.033	0.980± 0.033	0.958 ± 0.069	0.978 ± 0.050	0.940 ± 0.089	0.980 ± 0.018
mRMR	P1–P3,P5, P10, P11– P13 P16–P18	0.935 ± 0.045	0.934 ± 0.046	0.864 ± 0.0167	0.889 ± 0.106	0.860 ± 0.167	0.960± 0.043
ReliefF	P1–P4, P8–P10, P13, P16, P17, P19	0.985 ± 0.014	0.985 ± 0.014	0.970 ± 0.027	0.964 ± 0.050	0.980 ± 0.045	0.987 ± 0.018

5. Conclusions

This work developed and validated a compact multicondition identification framework for DHN transient pressure-wave signals, aiming to reliably distinguish leakage from common operation-induced disturbances under small-sample constraints. The main conclusions are as follows:

1. Integrated experimental validation under multi-condition DHN scenarios. A controllable indoor DHN testbed was established and a balanced four-class dataset (leakage, valve regulation, pump speed regulation, and emergency valve shut-off) was constructed, enabling systematic evaluation of leakage recognition against operation-induced transients.

2. **A robust and interpretable feature pipeline was formed by preprocessing, handcrafted features, and multiclass ReliefF selection.** Hampel filtering and a zero-phase 0.5–300 Hz band-pass filter were adopted to suppress impulsive spikes and broadband noise while preserving informative transient components. Time-/frequency-domain statistical descriptors were then extracted and screened by multiclass ReliefF to obtain a compact subset that retains discriminative information while reducing redundancy, which was further supported by feature-selection comparisons.

3. Feature-space, class-wise GAN generation effectively alleviated the small-sample limitation for training. By generating synthetic feature vectors in the selected feature space (training only), the framework enriched the training distribution without contaminating real-test evaluation, and the augmentation ratio was shown to have a non-monotonic effect—highlighting the necessity of tuning augmentation level rather than assuming "more is always better."

4. Improved multiclass identification and leakage-oriented performance on real test samples. Under outer 5-fold cross-validation on real data, GAN-generated feature augmentation improved generalization for representative classifiers. In particular, for random forest, accuracy and Macro-F1 increased from 0.960 to 0.985, while leakage recall rose from 0.920 to 0.980, indicating enhanced robustness when leakage signatures are partially confounded with operational disturbances.

Overall, the proposed framework provides a practical route toward fast and online-applicable DHN leakage recognition by combining robust transient-feature representation, classifier-independent feature screening, and feature-space generative augmentation under a strict train/test separation protocol.

 doi:10.20944/preprints202603.1448.v1

References

1. L. Manservigi, H. Bahlawan, E. Losi, M. Morini, P.R. Spina, M. Venturini. A diagnostic approach for fault detection and identification in district heating networks. Energy. 251 (2022) 14.
2. H. Bahlawan, N. Ferraro, A. Gambarotta, E. Losi, L. Manservigi, M. Morini, et al. Detection and identification of faults in a District Heating Network. Energy Conv Manag. 266 (2022) 15.
3. C. Liu, S.J. Zhou, Y.L. Zhang, C. Zhang, X.R. Liu. Leakage diagnosis of district heating-network based on system simulation and PCA_BP neural network. Process Saf Environ Protect. 180 (2023) 260-73.
4. E. Vollmer, J. Ruck, R. Volk, F. Schultmann. Leak detection using thermal imagery: Deep learning versus traditional computer vision state-of-the-art. ISPRS-J Photogramm Remote Sens. 228 (2025) 505-18.
5. Y. Koolivand, W.J. Szczerek, M.K. Dahl, J.M.H. Bendtsen, Y. Rezaeiyan, S. Richter, et al. Thermally Powered Autonomous Water Leakage Detection System for District Heating Pipes. IEEE Sens J. 25 (2025) 32953-63.
6. H.J. Li, H.H. Zhu, D.Y. Tan, B. Shi, J.H. Yin. Detecting pipeline leakage using active distributed temperature Sensing: Theoretical modeling and experimental verification. Tunn Undergr Space Technol. 135 (2023) 16.
7. M. Valincius, M. Vaisnoras, A. Kaliatka. Study and demonstration of pressure wave-based leak detection in a district heating network. Struct Infrastruct Eng. 14 (2018) 151-62.
8. X.J. Zheng, F.S. Hu, Y.R. Wang, L.J. Zheng, X.Y. Gao, H. Zhang, et al. Leak detection of long-distance district heating pipeline: A hydraulic transient model-based approach. Energy. 237 (2021) 16.
9. Z.Q. Xu, C. Li, L.B. Mu, S.L. Wang, J.H. Lu, Y.C. Lan. Leakage detection method of underground heating pipeline based on improved wavelet threshold function. Energy. 295 (2024) 16.
10. L.K. Zhao, Z. Cao, J.Q. Deng. A review of leak detection methods based on pressure waves in gas pipelines. Measurement. 236 (2024) 22.
11. X.J. Guo, J.Q. Deng, Z. Cao, Z.K. Xiong. Research on the generation and modulation of active pressure wave for pipelines leak diagnosis. Measurement. 242 (2025) 14.
12. X. Zhou, F. Liu, L.S. Luo, S.M. Peng, J.L. Xie. Study on the Leakage Diagnosis of a Chilled Water Pipeline Network System Based on Pressure Variation Rate Analysis for Climate Change Mitigation. Buildings-Basel. 15 (2025) 19.
13. K. Vahldiek, B. Rüger, F. Klawonn. Optimal Sensor Placement and Influence of Noise on Pressure Wave Evaluation for Leakage Localization in a District Heating Network. Sustain Energy Grids Netw. 30 (2022) 14.
14. E. Losi, L. Manservigi, P.R. Spina, M. Venturini. Data-driven approach for the detection of faults in district heating networks. Sustain Energy Grids Netw. 38 (2024) 14.
15. S.J. Zhou, Z. O'Neill, C. O'Neill. A review of leakage detection methods for district heating networks. Appl Therm Eng. 137 (2018) 567-74.
16. Z.K. Hu, B.Q. Chen, W.L. Chen, D.B. Tan, D.T. Shen. Review of model-based and data-driven approaches for leak detection and location in water distribution systems. Water Supply. 21 (2021) 3282-306.
17. J. Tang, X.F. Wang, H.B. Bi, R. Dai, Y.C. Yue, J.H. Yang, et al. Technology for detecting small-aperture leaks in natural gas pipelines utilizing transfer learning methodologies. Meas Sci Technol. 36 (2025) 14.
18. G.X. Shi, X.P. Wang, J.J. Zhang, X.L. Gao. High Precision Detection Pipe Bursts Based on Small Sample Diagnostic Method. Sensors. 25 (2025) 22.
19. D.M. Wang, Y. Sun, J.Y. Lu. Pipeline leak detection based on generative adversarial networks under small samples. Flow Meas Instrum. 101 (2025) 11.
20. M.M. Rajabi, P. Komeilian, X. Wan, R. Farmani. Leak detection and localization in water distribution networks using conditional deep convolutional generative adversarial networks. Water Res. 238 (2023) 14.
21. H.G. Zhang, X.G. Hu, D.Z. Ma, R. Wang, X.P. Xie. Insufficient Data Generative Model for Pipeline Network Leak Detection Using Generative Adversarial Networks. IEEE T Cybern. 52 (2022) 7107-20.
22. R. Shang, H.L. Dong, C. Wang, S.Q. Chen, T. Sun, C. Guan. Imbalanced data augmentation for pipeline fault diagnosis: A multi-generator switching adversarial network. Control Eng Practice. 144 (2024) 11.
23. G. Chandrashekar, F. Sahin. A survey on feature selection methods. Comput Electr Eng. 40 (2014) 16-28.
24. H.C. Peng, F.H. Long, C. Ding. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 27 (2005) 1226-38.

 doi:10.20944/preprints202603.1448.v1

25. R. Tibshirani. Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B-Stat Methodol. 58 (1996) 267-88.
26. J.C. Hu, S. Szymczak. A review on longitudinal data analysis with random forest. Brief Bioinform. 24 (2023) 11.
27. R.J. Urbanowicz, M. Meeker, W. La Cava, R.S. Olson, J.H. Moore. Relief-based feature selection: Introduction and review. J Biomed Inform. 85 (2018) 189-203.
28. N. Aggarwal, U. Shukla, G.J. Saxena, M. Rawat, A.S. Bafila, S. Singh, et al. Mean based relief: An improved feature selectionmethod based on ReliefF. Appl Intell. 53 (2023) 23004-28.
29. M. Robnik-Sikonja, I. Kononenko. Theoretical and empirical analysis of ReliefF and RReliefF. Mach Learn. 53 (2003) 23-69.
30. L.P. Liang, J. Zhang, K.J. Xu, G.Y. Ye, S.L. Yang, X.L. Yu. Classification modeling of valve internal leakage acoustic emission signals based on optimal wavelet scattering coefficients. Measurement. 236 (2024) 15.
31. L.Y. Wu, W. Liang, D.L. Sha. Cross-domain feature selection and diagnosis of oil and gas pipeline defects based on transfer learning. Eng Fail Anal. 143 (2023) 16.
32. A.F. Colombo, P. Lee, B.W. Karney. A selective literature review of transient-based leak detection methods. J Hydro-environ Res. 2 (2009) 212-27.
33. Z.Y. Xu, H.X. Liu, G.T. Fu, Y.K. Zeng, Y.C. Li. Feature selection of acoustic signals for leak detection in water pipelines. Tunn Undergr Space Technol. 152 (2024) 11.
34. M. Liang, T.F. An, R.H. Zhao, T.X. Liu, W.L. Liu. Signal processing techniques for detecting leakage in urban water supply pipelines: Denoising and feature enhancement. Tunn Undergr Space Technol. 162 (2025) 15.
35. L.H. Huang, B. Hu, S.T. Wan, B. Lu. Research on pipeline flange leakage detection method based on random forest and Pearson correlation coefficient. Appl Acoust. 240 (2025) 11.
36. Z. Ahmad, T.K. Nguyen, A. Rai, J.M. Kim. Industrial fluid pipeline leak detection and localization based on a multiscale Mann-Whitney test and acoustic emission event tracking. Mech Syst Signal Proc. 189 (2023) 15.
37. R.S. Liu, T. Zayed, R. Xiao. Advanced acoustic leak detection in water distribution networks using integrated generative model. Water Res. 254 (2024) 15.
38. R. Flórez-López. Reviewing RELIEF and its extensions:: A new approach for estimating attributes considering high-correlated features. 2nd IEEE International Conference on Data Mining. Ieee Computer Soc, Maebashi City, Japan, 2002. pp. 605-8.
39. H.Y. Fan, L.Y. Xue, Y. Song, M. Li. A repetitive feature selection method based on improved ReliefF for missing data. Appl Intell. 52 (2022) 16265-80.
40. J. Friedman, T. Hastie, R. Tibshirani. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 33 (2010) 1-22.
41. C. Cortes, V. Vapnik. SUPPORT-VECTOR NETWORKS. Mach Learn. 20 (1995) 273-97.
42. J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, A. Lopez. A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing. 408 (2020) 189-215.
43. X.D. Fan, X.W. Wang, X.J. Zhang, X. Yu. Machine learning based water pipe failure prediction: The effects of engineering, geology, climate and socio-economic factors. Reliab Eng Syst Saf. 219 (2022) 15.
44. K. Hattori, M. Takahashi. A new nearest-neighbor rule in the pattern classification problem. Pattern Recognit. 32 (1999) 425-32.
45. K. Taunk, S. De, S. Verma, A. Swetapadma, Ieee. A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. International Conference on Intelligent Computing and Control Systems (ICCS). Ieee, Anna Univ, Vaigai Coll Engn, Madurai, INDIA, 2019. pp. 1255-60.
46. J. Opitz. A Closer Look at Classification Evaluation Metrics and a Critical Reflection of Common Evaluation Practice. Trans Assoc Comput Linguist. 12 (2024) 820-36.
47. O. Rainio, J. Teuho, R. Klén. Evaluation metrics and statistical tests for machine learning. Sci Rep. 14 (2024) 14.

 doi:10.20944/preprints202603.1448.v1