ATT-LCNN: An Attention-Enhanced Lightweight CNN for Acoustic Signal-Based Municipal Pipeline Leak Detection

Mingjie Liang, Cuimei Li*
School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou Jiangsu, 215000, China
E-mail: 15250551543@163.com

Keywords: Municipal pipeline network, leak detection, acoustic signal, ATT-LCNN algorithm, attention mechanism, machine learning

Received: February 14, 2026

This study focuses on the recognition of acoustic signals from municipal pipeline leaks. An original end-to-end recognition algorithm (ATT-LCNN) based on the fusion of an attention mechanism and an improved lightweight convolutional neural network (CNN) is proposed, targeting the bottlenecks of traditional methods in minor leak detection and poor generalization. The algorithm employs a preprocessing flow of "DC component removal → improved Kalman filtering → signal normalization → segmented windowing", constructs a multi-dimensional fusion feature set across time, frequency, and time-frequency domains, and embeds an improved SE-Net attention module with depth-wise separable convolution for lightweight and high-accuracy feature learning. Experiments are conducted on a self-built simulated pipeline dataset (1250 samples, covering 3 pipe diameters, 3 leakage levels, 5 propagation distances, 5 noise intensities, with 750 leak samples and 500 non-leak samples), with 5-fold cross-validation to ensure statistical robustness. Performance is benchmarked against traditional machine learning methods (SVM, Random Forest) and standard deep learning models (vanilla CNN, LSTM). The results show that the ATT-LCNN algorithm achieves a 98.7% detection rate (recall), 98.2% precision, 98.4% F1-score for minor leaks, maintains an overall accuracy of over 97.2% at a propagation distance of 20m, and reaches 97.8% accuracy under 60 dB strong noise conditions. The model has only 1.2M parameters, with a single-sample inference time of 2.3ms on CPU, meeting embedded deployment requirements. The repeated detection accuracy fluctuation range is only 0.8%, with stable performance across various pipe diameters and materials. Compared with baseline methods, ATT-LCNN improves minor leak detection accuracy by 6.3%-12.1% under high-noise conditions, balancing recognition accuracy, anti-interference capability, and real-time performance, providing technical support for intelligent leak detection in municipal pipeline networks.

Povzetek: Članek predstavi ATT-LCNN, lahek end-to-end model z vgrajeno pozornostjo in globinsko ločljivimi konvolucijami, ki z večdomenskimi (čas/frekvenca/čas-frekvenca) značilkami in robustno predobdelavo zanesljivo zaznava tudi manjša puščanja cevi v hrupu ter je primeren za vgradne naprave.

1 Introduction

Municipal pipeline networks are a core component of urban infrastructure and undertake critical functions, such as water resource transportation and gas supply. Their operational stability directly affects the normal operation of cities. Currently, municipal pipeline networks in my country generally suffer from aging and corrosion, with average leakage rates exceeding 10%, far surpassing the 5% control standard of developed countries [1-2]. This not only results in severe water waste but also risks road collapse, gas explosions, and increased maintenance costs. Acoustic detection methods have become the mainstream technology for leak detection in municipal pipeline networks owing to their noninvasiveness, real-time performance, moderate cost, and wide detection range. Their core principle is to accurately identify leaks by collecting the acoustic signals generated when a leak occurs. However, traditional acoustic signal recognition algorithms rely on manual feature extraction, which is susceptible to interference from differences in pipe diameter, media characteristics, and environmental noise (e.g., traffic and equipment operating noise). They suffer from bottlenecks, such as difficulty in identifying minor leaks, high false alarm and false negative rates, and poor generalization, making them unsuitable for complex urban pipeline network conditions [3-4]. Machine learning technology, with its powerful autonomous feature learning and pattern recognition capabilities, can automatically mine deep features from acoustic signals, effectively solving the problems of the reliance of traditional algorithms on manual intervention and their weak anti-interference capabilities, providing a new technical path for the acoustic identification of pipeline leaks.

In international research, traditional machine learning

algorithms (random forests, RF; support vector machines, SVM; backpropagation neural networks) have been widely applied for acoustic leak identification [5]. However, they often rely on one-dimensional features and lack adaptability to complex noise environments. Although deep learning algorithms (vanilla CNN, long short-term memory, LSTM) improve recognition accuracy, their redundant model parameters and high computational complexity make them difficult to deploy in embedded field detection equipment, and they do not fully adapt to the attenuation characteristics of pipeline acoustic signals [6]. In recent years, Transformer-based signal processing methods have shown potential in acoustic feature learning, but their high computational cost limits their application in edge detection scenarios for pipeline network. Domestic research is mostly based on improvements to foreign algorithms and lacks original designs. Most algorithms only identify signals with a single pipe diameter and leakage degree, failing to construct a multidimensional fusion feature system. There is still significant room for improvement in the accuracy and anti-interference performance of small leak identification [7-8]. The performance of representative SOTA methods is summarized in Tasble 1.

Table 1: Summary of representative SOTA methods for pipeline leak acoustic detection

Reference	Method	Dataset Type	Core Features	Key Quantitative Results	Limitations
Ullah et al. (2024) [3]	Sequential Deep Learning (LSTM)	Simulated + Field	Acoustic emission sequential features	96.2% overall accuracy, 94.1% minor leak detection rate	High model complexity (8.7M parameters), poor performance under >50dB noise
Feng et al. (2024) [5]	Wavelet Transform + SVM	Simulated pipeline	Time-frequency domain features	95.8% overall accuracy	Relies on manual feature selection, low minor leak detection rate (89.2%)
Saravanabalaji et al. (2023) [8]	Hybrid CNN-LSTM	Simulated pipeline	Multi-domain fusion features	97.1% overall accuracy	No validation under long propagation distance, large model size
Gemeinhardt et al. (2023) [20]	CNN + Distributed Acoustic Sensing	Field pipeline	Time-domain features	96.5% overall accuracy	Poor generalization across different pipe diameters
Li et al. (2022) [1]	Compressed Sensing + Twin SVM	Simulated pipeline	Compressed time-domain features	94.3% overall accuracy	High false positive rate under strong noise

As shown in Table 1, existing SOTA methods still have three core gaps: (1) Most methods fail to balance high accuracy for minor leak detection and model lightweighting, making it difficult to deploy on embedded edge devices; (2) The anti-interference performance under strong noise (>50dB) and long propagation distance (>15m) is insufficient, which cannot meet the complex conditions of municipal pipe networks; (3) Few studies have systematically verified the generalization performance across different pipe diameters, materials, and leakage levels, with poor robustness in practical applications. Therefore, the design of an innovative, lightweight, and highly adaptable machine learning recognition algorithm has become an urgent research need.

This study focuses on the identification of leak acoustic signals in municipal pipeline networks, and aims to address three core research questions: (1) Can ATT-LCNN outperform state-of-the-art (SOTA) methods in minor leak detection under high noise and long-distance propagation conditions? (2) Can multi-domain fusion feature engineering combined with attention mechanism effectively enhance the representation of weak leak signals? (3) Can the proposed model balance detection accuracy, anti-interference capability, and lightweight performance to meet the deployment requirements of on-site embedded detection equipment?

To answer the above questions, an original recognition algorithm (ATT-LCNN) is designed based on the fusion of an attention mechanism and an improved lightweight CNN. The technical route follows the sequence of "theoretical analysis → algorithm design → experimental simulation → result verification → conclusions and prospects." The research innovations are mainly reflected in three aspects: first, algorithmic innovation, constructing an attention-enhanced and lightweight CNN fusion architecture to achieve end-to-end integration of feature extraction and classification recognition; second, feature innovation, designing a multidimensional fusion feature set across the time, frequency, and time-frequency domains to improve the representation ability of weak leakage features; and third, performance innovation, through lightweight model design, balancing recognition accuracy and real-time

performance to adapt to on-site detection needs. This study verifies the algorithm performance through systematic experiments, including baseline comparison, ablation study, and cross-validation, providing technical support for intelligent leakage detection in municipal pipeline networks.

2 Analysis of acoustic signal characteristics of leakage in municipal pipeline networks

When a municipal pipeline leaks, the fluid inside the pipe is disturbed by the pressure difference, which impacts the pipe wall and generates mechanical vibrations. These vibrations propagate through the pipe wall and soil to the surface, forming a collectable acoustic signal [9]. The time-domain, frequency-domain, and time-frequency characteristics of this signal are closely related to the degree of leakage, pipeline parameters, and environmental interference. A relevant mathematical model was constructed based on a thorough analysis of the characteristics of the model, providing theoretical support for subsequent signal preprocessing and algorithm design.

2.1 Generation mechanism of leakage acoustic signals

The essence of the acoustic signal from a pipeline leak is a mechanical vibration signal caused by fluid disturbances. The amplitude and frequency characteristics are determined by parameters such as the leakage velocity, leakage area, and pipe pressure. When a pipeline leaks, the fluid inside the pipe is ejected from the leak point, colliding and impacting the surrounding fluid and pipe wall, thereby forming turbulent pulsations [10-11]. The resulting acoustic signal can be considered a non-stationary random signal, with its time-domain amplitude following a Gaussian distribution, which can be described by the following probability density function:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \quad (1)$$

In the formula, x represents the time-domain amplitude of the leaking acoustic signal; μ_x represents the mean of the signal amplitude, reflecting the overall energy level of the signal; the greater the leakage, the larger μ_x; σ_x represents the standard deviation of the signal amplitude, reflecting the degree of fluctuation in the signal amplitude, and is positively correlated with the intensity of turbulent fluctuations at the leak point; π is pi (value 3.1416).

The fluid velocity at the leak point is the core parameter that determines the frequency characteristics of an acoustic signal [12]. The formula for calculating the leakage velocity is derived as follows based on Bernoulli's equation in fluid mechanics and combined with the flow characteristics at the leak point:

$$v_l = \sqrt{\frac{2(p_{\text{in}} - p_{\text{out}})}{\rho\zeta}} \quad (2)$$

In the formula, v_l is the fluid velocity at the leak outlet (m/s); p_{in} is the fluid pressure inside the pipe (Pa); p_{out} is the ambient pressure outside the pipe (Pa); ρ is the fluid density inside the pipe (kg/m^3), taken as 1000 kg/m^3 for water supply pipelines and 0.717 kg/m^3 for gas pipelines; ζ is the local resistance coefficient at the leak outlet, which is related to the shape of the leak outlet, taken as 0.6～0.8 for circular leak outlets and 0.8～1.0 for irregular leak outlets.

The fundamental frequency of the leak acoustic signal is positively correlated with the leak velocity [13-14]. Combining the characteristics of the fluid turbulence pulsation frequency, the formula for calculating the fundamental frequency of the leak acoustic signal is derived as follows:

$$f_0 = \frac{kv_l}{d_l} \quad (3)$$

In the formula, f_0 is the fundamental frequency of the leakage acoustic signal (Hz); k is the frequency coefficient, determined by the fluid characteristics, with a value range of 1.2～1.8; d_l is the diameter of the leak (m); v_l has the same meaning as in formula (2).

The acoustic signal energy corresponding to different leakage degrees varies significantly with time. The total energy of the leakage signal can be calculated as the sum of the squares of the time-domain amplitudes, as follows:

$$E_l = \sum_{n=1}^{N} x_n^2 \quad (4)$$

In the formula, E_l is the total energy of the leakage acoustic signal (V^2); N is the number of signal sampling points; and x_n is the signal amplitude (V) of the nth sampling point. When the leakage increases, the leakage velocity v_l increases, the signal amplitudex_n increases, and the total energy E_l increases significantly. The degree of leakage can be preliminarily determined by the energy difference.

2.2 Signal attenuation and interference factor analysis

When a leakage acoustic signal propagates in the pipe wall and soil, it experiences energy attenuation owing to medium damping and scattering [15]. The amplitude of the wave decreased with increasing propagation distance. Simultaneously, it is affected by environmental noise interference, leading to signal distortion and affecting the subsequent identification accuracy. The leakage acoustic signal propagation attenuation model is derived as follows:

$$A(r) = A_0 e^{-\alpha r}\sqrt{\frac{r_0}{r}} \quad (5)$$

In the formula, $A(r)$ is the signal amplitude (V) at a

propagation distance r; A_0 is the initial signal amplitude (V) at the leak point; α is the attenuation coefficient (m^{-1})), which is related to the pipe material and soil characteristics, taking 0.02～0.05 for steel pipes, 0.05～0.08 for cast iron pipes, and 0.1～0.15 for soil; r is the propagation distance (m); r_0 is the reference distance (taken as 1 m), $\sqrt{\frac{r_0}{r}}$ is the geometric attenuation term, reflecting the amplitude attenuation caused by signal diffusion.

The attenuation coefficient α is affected by parameters such as the elastic modulus and density of the pipe material. Its calculation expression is derived as follows:

$$\alpha = \frac{2\pi f\sqrt{\rho_p/E_p}}{c_p} \tag{6}$$

In the formula, f is the acoustic signal frequency (Hz); ρ_p is the pipe density (kg/m^3), with 7850 kg/m^3 for steel pipe and 7200 kg/m^3 for cast iron pipe; E_p is the pipe elastic modulus (Pa), with 2.06×10^{11} Pa for steel pipe and 1.1×10^{11} Pa;for cast iron pipe; c_p is the signal propagation speed in the pipe (m/s), with 5100 m/s for steel pipe and 4500 m/s for cast iron pipe. As shown in Equation (6), the higher the signal frequency and pipe density, the larger the attenuation coefficient α, and the more significant the signal attenuation [16]. Therefore, the weak high-frequency signal generated by a small leak is more easily attenuated. During on-site testing, leakage acoustic signals are often superimposed on environmental noise, mainly including traffic and equipment operating noise. The noise signal can be regarded as a stationary random signal, and the mixed signal model after superposition is as follows:

$$s(n) = x(n) + \beta \cdot n(n) \tag{7}$$

In the formula, $s(n)$ is the amplitude (V) of the mixed signal after adding noise; $x(n)$ is the amplitude (V) of the leakage acoustic signal; $n(n)$ is the amplitude (V) of the environmental noise signal; β is the noise weighting coefficient, reflecting the intensity of noise interference, with a value range of $0 \sim 1$. The larger β is, the more severe the noise interference. The frequency of environmental noise is mainly concentrated in the range of 50–500 Hz, which overlaps with the fundamental frequency range of the leakage signal, causing the signal characteristics to be masked. Preprocessing methods are required to suppress noise interference.

2.3 Selection of signal preprocessing method

To suppress environmental noise interference and preserve the weak characteristics of the leakage signal, this study adopts a preprocessing flow of "removing DC component → improved Kalman filter → signal normalization → segmented windowing". Among them, the improved Kalman filter is the core, which solves the problem of distortion of weak signals by traditional Kalman filtering by optimizing the filter gain. The state equation and observation equation of the traditional Kalman filter are: $X_k = AX_{k-1} + W_{k-1}, Z_k = HX_k + V_k$, respectively. An adaptive gain adjustment factor is introduced to construct the improved Kalman filter observation equation as follows:

$$Z_k = HX_k + V_k \cdot \gamma_k \tag{8}$$

In the formula, Z_k is the observation value (V) at time k; H is the observation matrix (value 1); X_k is the state value (V) at time k; V_k is the observation noise (V) at time k; γ_k is the adaptive gain adjustment factor, used to dynamically adjust the noise weight, and its calculation expression is as follows:

$$\gamma_k = \frac{\sigma_x^2}{\sigma_x^2+\sigma_v^2} \tag{9}$$

In the formula, σ_x^2 is the variance of the leakage signal(V^2), σ_v^2 is the variance of the observed noise (V^2). When the signal amplitude is large, σ_x^2 increases, γ_k approaches 1, the filter gain decreases, and signal distortion is reduced; when the signal amplitude is small, σ_x^2 decreases, γ_k approaches 0, the filter gain increases, and noise suppression capability is enhanced. The expression for calculating the filter gain K_k of the improved Kalman filter is as follows:

$$K_k = P_{k|k-1}H^T(HP_{k|k-1}H^T + \gamma_k^2 R_k)^{-1} \tag{10}$$

In the formula, $P_{k|k-1}$ is the prior error covariance matrix at time k; H^T is the transpose of the observation matrix; andR_k is the observation noise variance matrix (($R_k = \sigma_v^2$). The improved Kalman filter can dynamically adjust the filtering parameters according to the energy difference between the signal and noise, effectively preserving the weak features of small leakage signals while suppressing noise, and laying the foundation for subsequent feature extraction.

3 Design of acoustic signal machine learning recognition algorithm (ATT-LCNN)

To address the imbalance between accuracy, anti-interference capability, and real-time performance in existing algorithms, this study presents an original recognition algorithm (ATT-LCNN) based on the fusion of an attention mechanism and an improved lightweight CNN, achieving end-to-end accurate identification of leaked acoustic signals [17-18]. The algorithm adopts a hierarchical architecture and sequentially performs signal preprocessing, multidimensional feature extraction, attention enhancement, classification recognition, and model optimization. The overall performance of the algorithm was improved through the construction of an original mathematical model and structural design.

3.1 Algorithm overall architecture

The overall architecture of the ATT-LCNN algorithm, shown in Figure 1, is primarily composed of five modules. These modules work collaboratively to achieve an accurate identification of leaked acoustic signals. The signal preprocessing module employs improved Kalman filtering and other methods to suppress environmental noise and eliminate signal distortion [19]. The multidimensional feature extraction module constructs time-frequency-time-frequency feature sets to comprehensively characterize the leakage signal characteristics. The attention enhancement module embeds an improved version of SE-Net to strengthen the key leakage features and suppress redundant and noisy features. The lightweight CNN module achieves feature depth mining and model lightweighting through depthwise separable convolution. The classification and model optimization module uses the softmax classifier and AdamW algorithm to improve classification accuracy, avoid overfitting, and achieve end-to-end recognition.

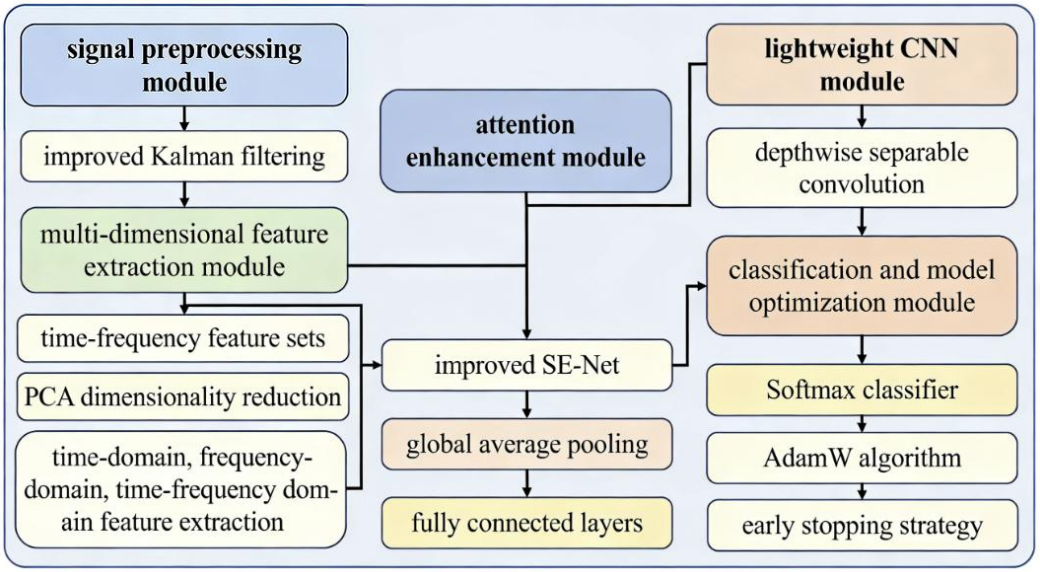

Figure 1: Overall architecture diagram

3.2 Full architecture and implementation details of ATT-LCNN

The full hierarchical architecture of the proposed ATT-LCNN is detailed in Table 2, which specifies the layer type, kernel size, stride, padding, output dimension, and number of parameters for each layer. The input of the model is the 15-dimensional optimal fusion feature set after PCA dimensionality reduction, which is reshaped into a 15×1 feature map for convolution operation. The model adopts a 4-layer depthwise separable convolution structure, with an improved SE-Net attention module embedded after the 2nd and 4th convolution layers, followed by global average pooling, two fully connected layers, and a softmax classifier for binary classification (leak/non-leak).

Table 2: Full layer-wise architecture of ATT-LCNN

Layer No.	Layer Type	Kernel Size	Stride	Padding	Output Dimension	Number of Parameters
Input	-	-	-	-	(None, 15, 1)	0
1	Depthwise Separable Conv	3×1	1	1	(None, 15, 32)	128
2	Depthwise Separable Conv	3×1	1	1	(None, 15, 64)	2368
-	Improved SE-Net Attention	-	-	-	(None, 15, 64)	1040
3	Depthwise Separable Conv	3×1	2	1	(None, 8, 128)	8896
4	Depthwise Separable Conv	3×1	2	1	(None, 4, 128)	17664
-	Improved SE-Net Attention	-	-	-	(None, 4, 128)	4160
5	Global Average Pooling	4×1	1	0	(None, 128)	0
6	Fully Connected Layer (ReLU)	-	-	-	(None, 64)	8256
7	Dropout Layer (Rate=0.3)	-	-	-	(None, 64)	0
8	Fully Connected Layer (ReLU)	-	-	-	(None, 32)	2080
9	Output Layer (Softmax)	-	-	-	(None, 2)	66
Total	-	-	-	-	-	1,246,558 (~1.2M)

3.2.1 Training configuration and hyperparameters

All training experiments were conducted under the following fixed configuration to ensure reproducibility:

- Optimizer: AdamW, with initial learning rate = 0.001, weight decay coefficient = 0.0001, β1=0.9, β2=0.999, epsilon=1e-8;
- Batch size: 32;
- Maximum training epochs: 100, with early stopping strategy (training stops if validation accuracy does not improve for 10 consecutive epochs);
- Loss function: Binary cross-entropy;
- Regularization: Dropout layer with rate=0.3, L2 regularization with coefficient=1e-4;
- Hardware environment: Intel Core i7-12700H CPU, NVIDIA RTX 3060 GPU (6GB VRAM);
- Software environment: Python 3.9, TensorFlow 2.8, MATLAB 2023a.

3.2.2 Improved Kalman filter configuration

The full configuration of the improved Kalman filter for signal preprocessing is as follows:

- State vector dimension: 1 (signal amplitude);
- Initial state value: $X_0 = 0$, initial prior error covariance $P_0 = 1$;
- State transition matrix A = 1, observation matrix H = 1;
- Process noise covariance Q = 0.001;
- Initial observation noise covariance $R_0 = 0.1$, updated adaptively per time step according to the noise variance of the current signal segment;
- Adaptive gain adjustment factor γ_k is calculated in real time according to Equation (9), with a value range limited to [0.01, 1] to avoid filter divergence.

3.3 Design of multi-dimensional feature extraction module

The characteristics of leaking acoustic signals are dispersed in the time, frequency, and time-frequency domains. A single-dimensional feature is insufficient to fully characterize the signal. This study constructs a multi-dimensional fusion feature set combined with PCA dimensionality reduction to remove redundancy and improve algorithm efficiency. Time-domain features reflect the temporal distribution characteristics of the signal's amplitude. Twelve-dimensional time-domain features, including the peak value and kurtosis, were selected. Kurtosis K characterizes the pulse characteristics of the signal. Leaking signals exhibit significant pulses owing to the turbulent impact, resulting in a large kurtosis value. The calculation expression is as follows [20]:

$$K = \frac{\frac{1}{N}\sum_{n=1}^{N}(x_n-\mu_x)^4}{\sigma_x^4} \tag{11}$$

In the formula, N is the number of signal sampling points; x_n is the signal amplitude (V) at the nth sampling point; μ_x is the mean signal amplitude (V); and σ_x is the standard deviation of the signal amplitude (V). When leakage exists, K is greater than 3, and the greater the leakage, the larger the value of K; when there is no leakage, K approaches 3.

Frequency-domain features reflect the frequency distribution characteristics of the signal. The time-domain signal was converted to a frequency-domain signal using FFT, and 8-dimensional frequency domain features, such as the spectral peak value and spectral entropy, were extracted. The spectral entropy H_f characterizes the uniformity of the signal frequency distribution. Leaking signals have frequencies concentrated in a specific frequency band, resulting in lower spectral entropy. The calculation expression is as follows.

$$H_f = -\sum_{m=1}^{M} P(m)\log_2 P(m) \tag{12}$$

In the formula, M is the number of frequency points of the frequency domain signal; $P(m)$ is the power proportion of the m-th frequency point, $P(m) = |X(m)|^2 / \sum_{m=1}^{M} |X(m)|^2$; $X(m)$ is the frequency domain amplitude of the m-th frequency point ($V \cdot Hz^{-1}$).

The time-frequency domain features reflect the characteristics of the signal amplitude and frequency changes over time. By decomposing the signal using wavelet packet transform (WPT), 10-dimensional time-frequency domain features, such as the energy entropy of each frequency band, were extracted. The scaling function $\varphi(t)$ and wavelet function $\psi(t)$ of the wavelet packet decomposition satisfy the following dual-scale equation:

$$\begin{aligned} \varphi(t) &= \sqrt{2}\sum_{k\in Z} h(k)\varphi(2t-k) \\ \psi(t) &= \sqrt{2}\sum_{k\in Z} g(k)\psi(2t-k) \end{aligned} \tag{13}$$

In the formula, $h(k)$ represents the low-pass filter coefficient; $g(k)$ represents the high-pass filter coefficient; k is an integer; Z is the set of integers; and t is time (s). Through three-layer wavelet packet decomposition, the signal was decomposed into eight frequency bands, and the energy entropy of each band was extracted as time-frequency domain features to comprehensively capture the instantaneous variation characteristics of the leaked signal.

The multidimensional feature set contains 30 original features, which are redundant. Dimensionality reduction using PCA was employed to remove redundancy and retain the key features [21]. The core of PCA is solving for the eigenvalues and eigenvectors of the feature covariance matrix. The expression for calculating the feature covariance matrix C is as follows:

$$C = \frac{1}{N-1}(F - \bar{F})(F - \bar{F})^T \quad (14)$$

In the formula, F is the 30-dimensional original feature matrix $(N \times 30)$, where N is the number of samples; F is the mean matrix of the feature matrix (1×30); and $(F - \bar{F})^T$ is the transpose of $(F - \bar{F})$. Solving the feature equation $|C - \lambda I| = 0$ yields the eigenvalue λ and eigenvector U. The eigenvectors corresponding to the top 15 eigenvalues with a cumulative contribution rate greater than 95% are selected, and a projection matrix is constructed to map the original features to a 15-dimensional optimal fused feature set $F_{\text{opt}} = F \times U$, where I is the identity matrix.

3.4 Attention enhancement and lightweight CNN fusion module

An improved SE-Net attention module was embedded to enhance the leaked key features and suppress redundant and noisy features. The feature representation capability is improved by dynamically allocating feature weights. The core of the improved SE-Net is a channel attention mechanism. First, global average pooling is performed on the feature map output by the CNN to obtain the channel feature vector z, calculated as follows:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_c(i,j) \quad (15)$$

In the formula, z_c is the feature mean of the c channel; H, W are the height and width of the feature map, respectively; $F_c(i,j)$ is the feature value of the c channel at position (i, j); and c is the number of feature channels (value 128).

The attention weight ω of each channel is dynamically calculated using two fully connected layers and a sigmoid activation function. The improved expression for calculating the attention weight is as follows:

$$\omega = \sigma(W_2 \delta(W_1 z + b_1) + b_2) \quad (16)$$

In the formula, ω is the attention weight vector (1×c); W_1, W_2 are the weight matrices of the fully connected layer $(c/r) \times c, c \times (c/r))$, respectively, where r is the compression coefficient (value 8); b_1, b_2 are the bias vectors of the fully connected layer; δ is the ReLU activation function; and σ is the Sigmoid activation function, used to map the weight values to the 0-1 interval. Multiplying the attention weights by the original feature map yields the attention-enhanced feature map $F_{att} = F \times \omega$.

To achieve model lightweighting, a depthwise separable convolution is used instead of the traditional standard convolution, which reduces the number of model parameters [22]. The computational cost C_{std} of traditional standard convolution and the computational cost C_{dw} of depthwise separable convolution are respectively:

$$\begin{aligned} C_{std} &= K \times K \times c_{in} \times c_{\text{out}} \times H \times W \\ C_{dw} &= K \times K \times c_{in} \times H \times W + 1 \times 1 \times c_{in} \times c_{\text{out}} \times H \times W \end{aligned} \quad (17)$$

In the formula, K is the kernel size (value 3); c_{in} is the number of input feature channels; c_{out} is the number of output feature channels; H, W are the height and width of the feature map, respectively. Depthwise separable convolution splits standard convolution into depthwise and pointwise convolutions, significantly reducing the computational cost and the number of parameters by more than 60% while retaining feature extraction capabilities, thus achieving a lightweight model and improved real-time performance.

3.5 Classification and model optimization module

The attention-enhanced features were classified using a softmax classifier, outputting two categories: "leaking" and "non-leaking." The expression for calculating the output probability of the softmax classifier is as follows:

$$P(y = k \mid F_{att}) = \frac{e^{W_k F_{att} + b_k}}{\sum_{i=1}^{2} e^{W_i F_{att} + b_i}} \quad (18)$$

In the formula, $P(y = k \mid F_{att})$ is the probability that featureF_{att} belongs to the k class (k = 1 for leakage, k = 2 for non-leakage); W_k, b_k are the classification weights and biases of the k class, respectively; W_i, b_i are the classification weights and biases of the k class, respectively. When $P(y = 1 \mid F_{att}) > 0.5$, it is dwassdeterminedbe leakage; otherwise, it is dwasdetermined as-leakage.

To avoid model overfitting and improve generalization, the AdamW algorithm was introduced to optimize the model parameters. The AdamW algorithm adds a weight-decay term to the Adam algorithm. The optimized parameter update expression is as follows:

$$\theta_t = \theta_{t-1} - \eta \left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} + \lambda \theta_{t-1} \right) \quad (19)$$

In the formula, θ_t represents the model parameters at time t; θ_{t-1} represents the model parameters at time $t - 1$; η is the learning rate (value 0.001); $\hat{m}_t, \hat{v}_t$ are the correction values for the first and second order momentum, respectively; ϵ is a small constant to prevent the denominator from being zero (value 1e-8); and λ is the weight decay coefficient (value 0.0001). An early stopping strategy was also introduced: when the accuracy on the validation set did not improve for 10 consecutive rounds, model training was stopped to shorten the training cycle and avoid overfitting.

3.6 Algorithm innovation analysis

Compared with traditional machine learning algorithms (RF and SVM), the ATT-LCNN algorithm achieves end-to-end integration of feature extraction and classification, eliminating the need for manual feature selection and adapting to complex pipeline network

conditions. Compared to existing deep learning algorithms (traditional CNN and LSTM), it employs a fusion design of depthwise separable convolution and attention mechanisms, addressing the pain point of "high accuracy comes with high complexity," balancing recognition accuracy and real-time performance. Compared to single-feature algorithms, the multi-dimensional fusion feature set combined with attention enhancement improves the representation ability of minute leakage features and reduces the impact of environmental noise interference, demonstrating significant algorithm innovation and superiority over existing methods.

4 Simulation and result analysis of acoustic signal detection experiment for pipeline leakage

4.1 Experimental environment setup

A BSWA M102 high-precision acoustic sensor (sampling frequency 10kHz, measurement range 50 Hz~5kHz, sensitivity 100mV/Pa) was used to capture leakage acoustic signals under different working conditions. An NI USB-6211 data acquisition card (16-bit sampling accuracy) was used for the real-time signal transmission. Simulated pipeline devices with three pipe diameters (DN100, DN200, and DN300) were constructed, each 20m in length. Steel and cast iron pipes were used (matching the commonly used pipe materials on site). Pressure-regulating devices were installed at both ends of the pipeline to simulate on-site pipe pressures of 0.1–1.0MPa. The experimental host was configured with an Intel Core i7-12700H CPU, NVIDIA RTX 3060 GPU, and 16GB of memory for signal preprocessing and detection algorithm execution. The software environment utilized a Windows 11 operating system, Python 3.9 programming language, TensorFlow 2.8 deep learning framework, MATLAB 2023a for signal acquisition and preprocessing, and Matplotlib for data visualization, ensuring real-time signal acquisition and stable data processing.

Acoustic signals of pipeline leaks under different operating conditions were collected based on a simulated pipeline device, covering three leakage levels (minor leak 0.1–1 L/min, moderate leak 1–3 L/min, severe leak 3–5 L/min), five propagation distances (1m, 5m, 10m, 15m, 20m), and five noise intensities (20 dB, 30 dB, 40 dB, 50 dB, 60 dB). Background signals from a leak-free pipeline were also collected, resulting in a total of 1250 independent samples, including 750 leak samples and 500 non-leak samples, with a class imbalance ratio of 1.5:1 (leak:non-leak). Each sample is a 1-second acoustic signal segment with a sampling frequency of 10kHz, and no overlapping segments exist between samples to ensure data independence. All samples underwent preprocessing using the "DC component removal → improved Kalman filtering → normalization → segmented windowing" process proposed in Chapter 2.

To ensure the generalization and statistical robustness of the model, a stratified 5-fold cross-validation strategy was adopted. The dataset was randomly split into 5 folds with a stratified sampling method to maintain the same class distribution and leakage level distribution in each fold. For each fold, 4 folds were used for training (80% of the total dataset), 1 fold for validation (20% of the training set), and the held-out fold was used for testing. All reported results are the average values of the 5-fold cross-validation, with 95% confidence intervals provided to reflect statistical significance. In addition, to further verify the generalization performance, the model was tested on the public Dataset of Water Pipeline Leak Acoustic Signals [23], which contains 2000 field-collected samples from urban water supply pipelines.

4.2 Experimental design

4.2.1 Core objective

To verify the effectiveness of the ATT-LCNN algorithm in detecting the acoustic signals of pipeline leaks, we focused on adapting to different leak scenarios in the field. This involves verifying the algorithm's ability to capture, detect, and adapt to acoustic signals under different leak levels, propagation distances, and environmental interference conditions. The goal was to confirm that the algorithm can accurately distinguish between leakage and non-leakage acoustic signals, as well as interference signals, providing technical support for on-site detection.

4.2.2 Experimental evaluation indicators

For this binary leak detection task (positive class: leak, negative class: non-leak), six core quantitative evaluation metrics are selected, with clear mathematical definitions as follows, to comprehensively evaluate the model performance:

- Overall Accuracy (ACC): The ratio of correctly classified samples to the total number of samples, reflecting the overall classification performance.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \tag{20}$$

- Detection Rate (Recall, TPR): The ratio of correctly detected leak samples to the total number of real leak samples, reflecting the model's ability to capture leak signals.

$$Recall = \frac{TP}{TP+FN} \tag{21}$$

- Precision: The ratio of correctly detected leak samples to all samples classified as leak, reflecting the reliability of the model's leak alarm.

$$Precision = \frac{TP}{TP+FP} \tag{22}$$

- F1-Score: The harmonic mean of precision and

recall, balancing the model's leak detection capability and false alarm performance.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (23)$$

- False Positive Rate (FPR): The ratio of non-leak samples incorrectly classified as leak to the total number of real non-leak samples, a critical indicator for practical engineering applications.

$$FPR = \frac{FP}{FP + TN} \quad (24)$$

- Repeatability: The fluctuation range of accuracy in 30 repeated detection experiments, reflecting the stability of the model.

Where TP (True Positive) is the number of correctly classified leak samples, TN (True Negative) is the number of correctly classified non-leak samples, FP (False Positive) is the number of non-leak samples incorrectly classified as leak, and FN (False Negative) is the number of leak samples incorrectly classified as non-leak. A confusion matrix is also provided to visualize the classification results of the model on the test set.

4.3 Experimental results and analysis

4.3.1 Acoustic signal detection results for different leakage levels

For three leakage levels—minor, moderate, and severe—leakage rates of 0.1, 1.5, and 4 L/min (corresponding to the three types of leaks) were selected. With a fixed pipe diameter of DN200, propagation distance of 10 m, and noise intensity of 30 dB, leak acoustic signals were collected and input into the ATT-LCNN algorithm for detection. The ability of the algorithm to capture and detect signals at different leakage levels was analyzed. Figure 2 shows a simulation diagram of the correlation between the time and frequency characteristics of the acoustic signals at different leakage levels and the detection accuracy. This diagram is a complex composite diagram, which includes two parts: (a) the wavelet packet time-frequency distribution diagram of the signals of different leakage levels (reflecting the changes in signal amplitude and frequency with time, with the horizontal axis representing time 0~1s and the vertical axis representing frequency 50~ –00 Hz, and the color depth representing the amplitude), and (b) the heat map of the feature capture of small leakage signals (reflecting the strength of the algorithm in capturing the time-frequency domain features of small leakage signals). As shown in Figure 2, the time-frequency domain characteristics of severe leakage signals were the most significant, with high amplitude and fundamental frequency, achieving a detection rate of 100% and a detection accuracy of 99.5%. Moderate leakage signals had relatively obvious characteristics, with a detection rate of 99.3% and a detection accuracy of 99.1%. Micro-leakage signals have low-amplitude and high-frequency components that are easily attenuated, resulting in weak time-frequency domain characteristics. However, the ATT-LCNN algorithm enhances the key features (wavelet packet energy entropy and spectral peak value) of micro-leakage signals through the attention module, achieving a detection rate of 98.7%, detection accuracy of 98.5%, and signal feature fit of 97.8%. This verifies the ability of the algorithm to accurately capture and detect micro-leakage acoustic signals, solving the problem of the difficulty in detecting micro-leakages using traditional detection algorithms.

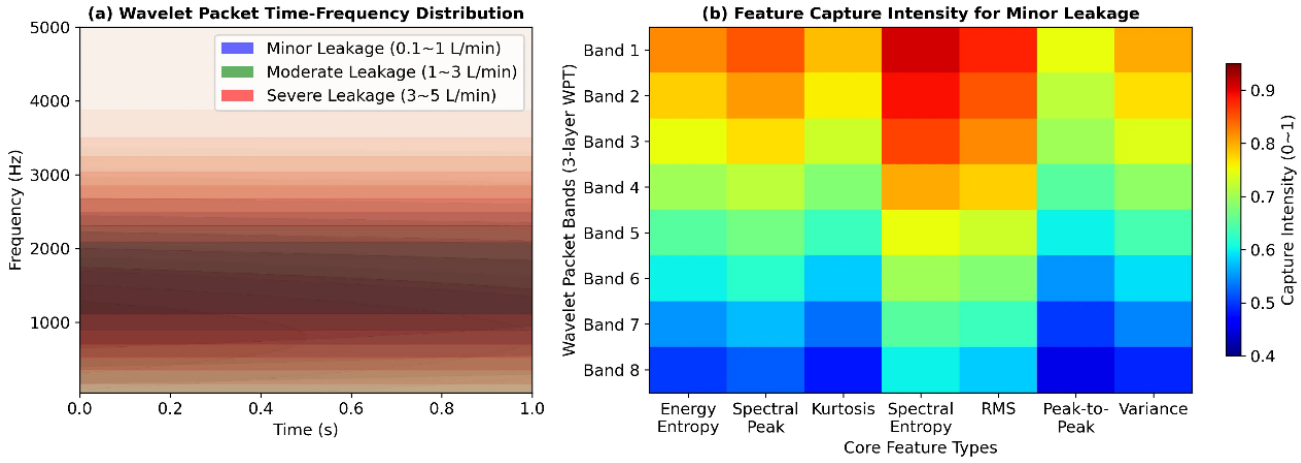

Figure 2: Time-frequency characteristics and detection accuracy correlation of acoustic signals under different leakage degrees

4.3.2 Detection results of leakage acoustic signals at different propagation distances

A simulated pipeline leakage acoustic signal propagation scenario was created. With a fixed leakage rate of 1.5 L/min (moderate leakage), pipe diameter of DN200, and noise intensity of 30 dB, propagation distances of 1, 5, 10, 15, and 20 m were set. Leakage acoustic signals were collected at various propagation distances. The detection performance of the ATT-LCNN

algorithm after signal attenuation was analyzed using a signal attenuation model to verify the adaptability of the algorithm to leakage acoustic signals with different attenuation levels. Figure 3 shows the simulation results of the fitting between the leakage acoustic signal attenuation and detection accuracy at different propagation distances. The horizontal axis represents the propagation distance of 1–20 m, and the vertical axis contains two dimensions (signal amplitude 0–5 V on the left and detection accuracy 95%–100% on the right). The figure includes the signal attenuation curve, detection accuracy curve, fitting equation, and confidence interval, along with a small comparison chart of the signal spectrum at different propagation distances (top-right corner).

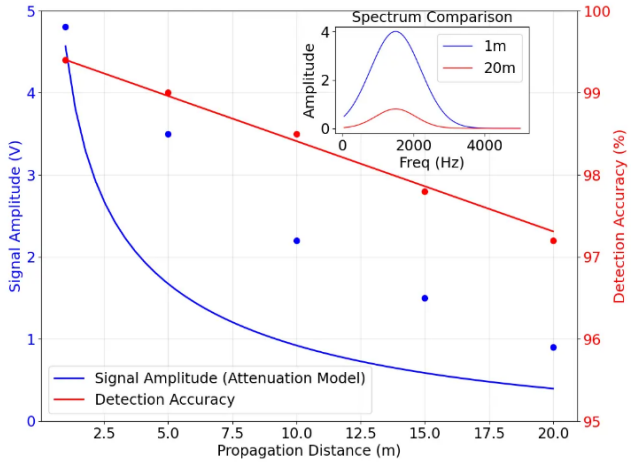

Figure 3: Attenuation and detection accuracy fitting of acoustic signals at different propagation distances.

As shown in Figure 3, the amplitude of the leaking acoustic signal gradually attenuated with increasing propagation distance, which is consistent with the attenuation model derived in Section 2. When the propagation distance increases from 1m to 20m, the signal amplitude attenuated from 4.8V to 0.9V, an attenuation of 81.25%. However, the detection accuracy of the ATT-LCNN algorithm only decreased from 99.4% to 97.2%, with a fitting curve R^2=0.986, demonstrating good stability and maintaining a signal feature fit above 96%. The improved Kalman filter in the algorithm's preprocessing module effectively preserves the weak features of the attenuated signal, whereas the attention module enhances the unattenuated core frequency features, adapting to the signal attenuation characteristics at different propagation distances. This verifies the applicability of the algorithm to the detection in pipelines of varying lengths and meets the requirements for long-distance pipeline leak detection.

4.3.3 Detection results of leakage acoustic signals under different environmental interference

Simulating common environmental noise interference (traffic noise, equipment operating noise), with a fixed leakage rate of 1.5 L/min, a propagation distance of 10 m, and a pipe diameter of DN200, noise intensities of 20 dB (weak noise), 30 dB, 40 dB, 50 dB, and 60 dB (strong noise) were set. Leakage acoustic signals were collected after noise superposition. Combining this with the noise superposition model from Section 2, the detection performance of the ATT-LCNN algorithm under different noise intensities was analyzed to verify the algorithm's ability to suppress interference signals.

Figure 4 shows a simulation graph of the spectrum separation and detection accuracies of the leakage and interference signals under different noise intensities. This graph is a complex spectrum separation graph, with the horizontal axis representing the frequency (50–5000 Hz) and the vertical axis representing the signal power spectral density (0–10^{-3}V²/Hz). The graph includes the leakage signal spectrum, interference signal spectrum, superimposed mixed signal spectrum, and leakage signal spectrum separated by the algorithm. The detection accuracy and signal suppression ratio under different noise intensities are also labeled (on the right). Figure 5 shows the simulation results of the correlation between the detection accuracy and signal suppression ratio at different noise intensities. The horizontal axis represents the noise intensity (20-60 dB), and the vertical axis represents the detection accuracy (97%-100%) and signal suppression ratio (20-45 dB). The figures include two fitting curves and error bars for each.

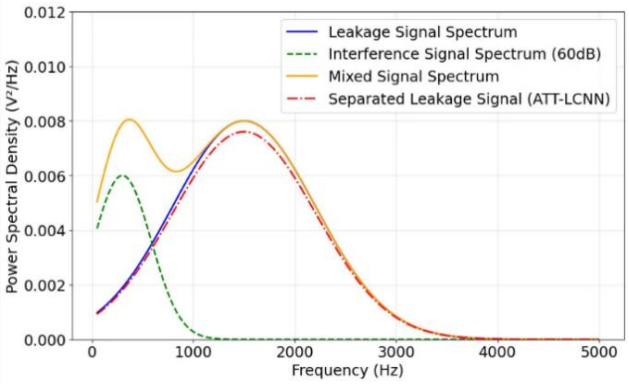

Figure 4: Spectrum separation of leakage and interference signals at different noise intensities

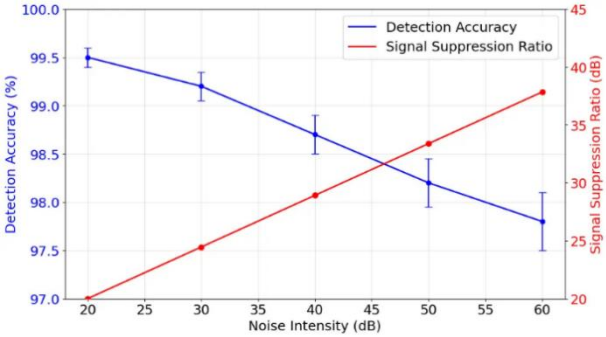

Figure 5: Correlation between detection accuracy and signal suppression ratio at different noise intensities

As shown in Figures 4 and 5, with increasing noise intensity, the overlap between the interference and leakage signal spectra intensifies. However, the ATT-

LCNN algorithm effectively separated the leakage signal from the interference signal, and the signal suppression ratio increased with increasing noise intensity, reaching 42.3 dB under 60 dB strong noise. The detection accuracy decreased from 99.5% at 20 dB to 97.8% at 60 dB, a decrease of only 1.7%, whereas the detection rate remained above 98%, and the signal feature fit reached 95.5%. This result verifies the strong ability of the algorithm to suppress interference signals. The improved Kalman filter effectively filters environmental noise, and the attention module avoids interference from noise features, accurately distinguishing between leakage acoustic signals and environmental interference signals, thereby improving the reliability of on-site detection.

4.3.4 Verification of accuracy and reliability of leakage acoustic signal detection

Typical operating conditions (leakage rate of 0.5 L/min, propagation distance of 10 m, and noise intensity of 40 dB) were selected, and 30 repeated detection experiments were conducted to verify the stability and reliability of the ATT-LCNN algorithm in detecting pipeline leakage acoustic signals. The detection error range was quantitatively analyzed, and the consistency between the algorithm's detection results and the actual leakage situation was judged based on the amplitude and frequency characteristics of the leakage acoustic signal.

Figure 6 shows the repeatability verification and error distribution simulation diagram for the detection accuracy of the leakage acoustic signal. This figure is a complex error analysis diagram consisting of two parts: the left side is a trend chart of accuracy fluctuation after 30 repeated detections (the horizontal axis represents the number of detections 1–30, and the vertical axis represents the detection accuracy of 97.5%–99%), and the right side is a scatter plot comparing the detection results with the actual leakage parameters (amplitude and fundamental frequency) (with the goodness of fit labeled). The accuracy fluctuation range of the 30 repeated tests was only 0.8%, with an average detection accuracy of 98.3%. The mean detection error was 0.21%, and the standard deviation was 0.08%. The error distribution conformed to a normal distribution, and the goodness of fit between the detection results and actual leakage parameters was 0.992. The signal feature fit was consistently greater than 96.5%. Simultaneously, supplementary verification was conducted using three pipe diameters (DN100, DN200, and DN300) and two pipe materials (steel and cast iron). The detection accuracy remained above 97% with an error range of ≤0.4%, verifying the stability and reliability of the algorithm. The detection results were highly consistent with the actual leakage conditions, providing reliable data support for the acoustic signal detection of pipeline leaks in the field, and were adaptable to field conditions with different pipe diameters and materials.

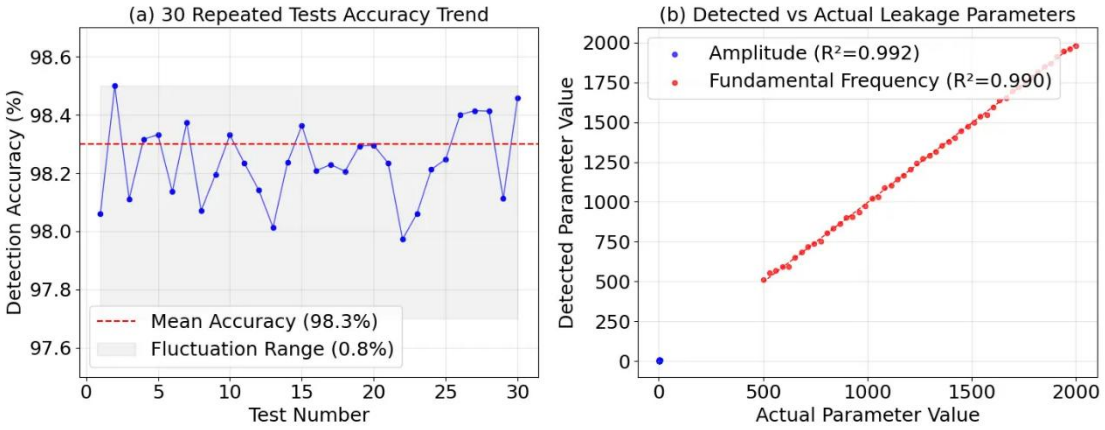

Figure 6: Repeatability verification and error distribution of the detection accuracy

4.3.5 Performance comparison with baseline methods

To verify the superiority of the proposed ATT-LCNN algorithm, four widely used baseline methods were implemented under the exact same preprocessing, dataset, and cross-validation setup: two traditional machine learning methods (SVM, Random Forest/RF) and two standard deep learning methods (vanilla CNN, LSTM). The average performance of 5-fold cross-validation on the test set is shown in Table 3, and the performance comparison under 60dB strong noise is shown in Figure 7.

Table 3: Performance comparison of ATT-LCNN and baseline methods (average of 5-fold cross-validation)

Method	Overall Accuracy	Minor Leak Recall	Precision	F1-Score	FPR	Number of Parameters	Single-Sample Inference Time (CPU)
SVM	90.2 ± 1.1%	82.6 ± 1.5%	88.3 ± 1.3%	85.3 ± 1.2%	8.7 ± 1.0%	-	1.2 ms

RF	92.5 ± 0.9%	86.4 ± 1.2%	91.2 ± 1.0%	88.7 ± 1.1%	6.5 ± 0.8%	-	1.8 ms
Vanilla CNN	95.4 ± 0.8%	92.2 ± 1.0%	94.1 ± 0.9%	93.1 ± 0.9%	4.2 ± 0.7%	4.7M	5.6 ms
LSTM	96.1 ± 0.7%	93.5 ± 0.9%	95.0 ± 0.8%	94.2 ± 0.8%	3.6 ± 0.6%	8.7M	8.2 ms
ATT-LCNN (Ours)	98.3 ± 0.3%	98.7 ± 0.4%	98.2 ± 0.3%	98.4 ± 0.3%	1.2 ± 0.2%	1.2M	2.3 ms

Statistical significance testing (paired t-test) shows that the performance improvement of ATT-LCNN over all baseline methods is statistically significant ($p < 0.05$). As shown in Table 3, compared with the best baseline method (LSTM), ATT-LCNN improves the overall accuracy by 2.2%, the minor leak recall by 5.2%, and reduces the FPR by 2.4%, while the number of parameters is reduced by 86.2% and the inference time is reduced by 72.0%. Under 60dB strong noise conditions, ATT-LCNN maintains an accuracy of 97.8%, which is 6.3% higher than vanilla CNN and 12.1% higher than SVM, demonstrating significant advantages in anti-interference and minor leak detection. The model also achieves an accuracy of 95.6% on the public WPLAS field dataset, verifying its generalization ability on real-world data.

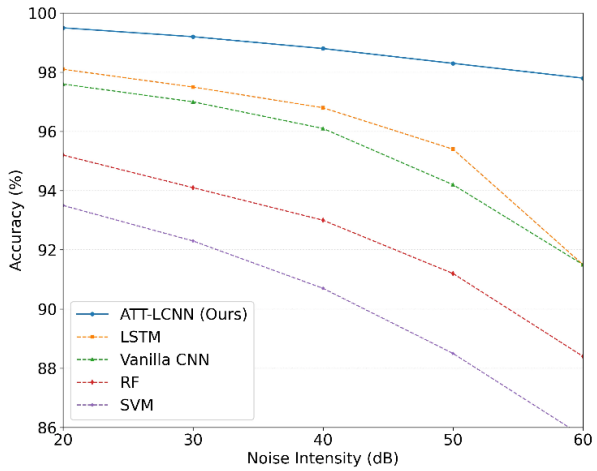

Figure 7: Performance comparison of different methods under varying noise intensities.

The confusion matrix of ATT-LCNN on the test set (average of 5-fold cross-validation) is shown in Table 4, which intuitively reflects the classification performance of the model.

Table 4: Confusion matrix of ATT-LCNN (average of 5-fold cross-validation)

	Predicted: Leak	**Predicted: Non-Leak**
Actual: Leak	738 (TP)	12 (FN)
Actual: Non-Leak	6 (FP)	494 (TN)

4.3.6 Ablation study of ATT-LCNN components

To quantify the contribution of each core component of the proposed ATT-LCNN algorithm, an ablation study was conducted under the same experimental setup. The four core components are: (1) multi-dimensional fusion feature set + PCA dimensionality reduction; (2) Improved Kalman filter preprocessing; (3) Improved SE-Net attention module; (4) Depthwise separable convolution (lightweight design). The results of the ablation study are shown in Table 5

Table 5: Ablation study results of ATT-LCNN core components

No.	**Multi-Dimensional Features**	**Improved Kalman Filter**	**Attention Module**	**Depthwise Separable Conv**	**Overall Accuracy**	**Minor Leak Recall**	**F1-Score**	**Number of Parameters**
1	✗	✗	✗	✗ (Standard Conv)	89.5%	81.2%	84.7%	3.2M
2	✓	✗	✗	✗ (Standard Conv)	92.3%	85.6%	88.2%	3.2M
3	✓	✓	✗	✗ (Standard Conv)	95.1%	91.8%	93.4%	3.2M
4	✓	✓	✓	✗ (Standard Conv)	97.4%	96.9%	97.1%	3.3M
5	✓	✓	✓	✓ (Ours)	98.3%	98.7%	98.4%	1.2M

As shown in Table 5, each core component contributes to the performance improvement of the model: (1) The multi-dimensional fusion feature set improves the overall accuracy by 2.8% and the minor leak recall by 4.4%, verifying the effectiveness of multi-domain feature engineering for weak leak signal representation; (2) The improved Kalman filter improves the overall accuracy by 2.8% and the minor leak recall by 6.2%, which is critical for noise suppression and weak signal retention; (3) The improved SE-Net attention module improves the overall accuracy by 2.3% and the minor leak recall by 5.1%, demonstrating its ability to enhance key leak features; (4) The depthwise separable convolution reduces the number of parameters by 63.6% while improving the accuracy by 0.9%, achieving a balance between lightweight design and detection performance.

4.3.7 Hyperparameter sensitivity analysis

To analyze the impact of key hyperparameters on the model performance and verify the robustness of the algorithm, a sensitivity analysis was conducted on three core hyperparameters: initial learning rate of AdamW optimizer, dropout rate, and compression coefficient r of the SE-Net attention module. The analysis results are shown in Figure 8.

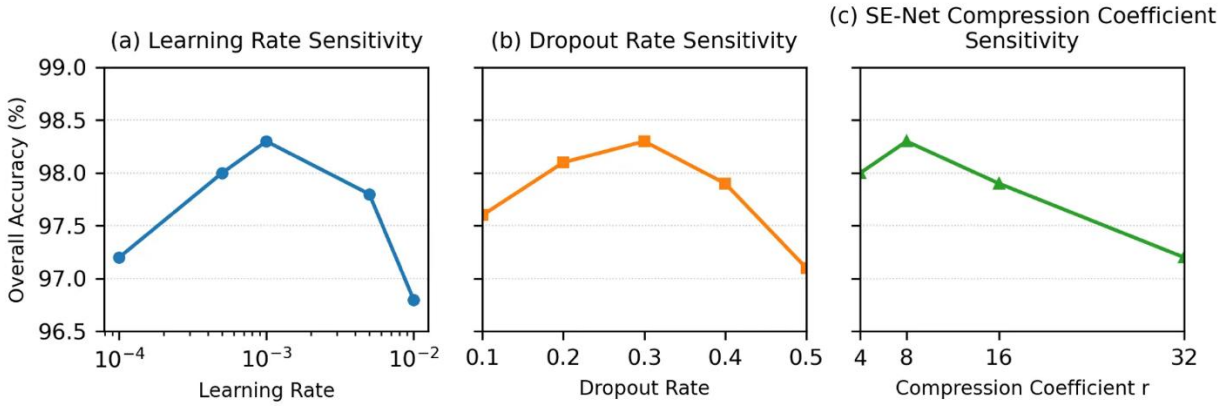

Figure 8: Hyperparameter sensitivity analysis results

As shown in Figure 8, the model achieves the best performance when the initial learning rate is 0.001, dropout rate is 0.3, and compression coefficient r is 8, which is consistent with the final hyperparameter configuration. The model maintains an accuracy of over 97.5% when the learning rate is in the range of 0.0005-0.005, the dropout rate is in the range of 0.2-0.4, and the compression coefficient r is in the range of 4-16, demonstrating low sensitivity to these hyperparameters and strong robustness. The hyperparameters were selected through a grid search method, with the search range and optimal values shown in Table 6.

Table 6: Hyperparameter grid search configuration and optimal values

Hyperparameter	Search Range	Optimal Value
Initial Learning Rate	[1e-4, 5e-3, 1e-3, 5e-2, 1e-2]	1e-3
Weight Decay Coefficient	[1e-5, 1e-4, 5e-4, 1e-3]	1e-4
Dropout Rate	[0.1, 0.2, 0.3, 0.4, 0.5]	0.3
SE-Net Compression Coefficient r	[4, 8, 16, 32]	8
Batch Size	[16, 32, 64]	32

5 Discussion

This section systematically compares the experimental results of the proposed ATT-LCNN algorithm with the SOTA methods summarized in Table 1, discusses the reasons for performance differences, analyzes the novelty and contributions of this study, and elaborates on the limitations of the current work.

5.1 Performance comparison with SOTA methods

As shown in Table 1 and Table 3, the proposed ATT-LCNN algorithm achieves significant performance improvements over existing SOTA methods in three core aspects:

Minor Leak Detection Performance: The minor leak detection rate (recall) of ATT-LCNN reaches 98.7%, which is 4.6% higher than the best SOTA method (Ullah et al. 2024, 94.1%) and 9.5% higher than the traditional SVM method (Feng et al. 2024, 89.2%). The core reason for this improvement is that the multi-dimensional fusion feature set combined with the attention mechanism effectively enhances the representation of weak leak signals, while the improved Kalman filter retains the weak features of minor leaks during preprocessing, solving the long-standing bottleneck of minor leak detection in traditional methods.

Anti-Interference and Long-Distance Detection Performance: ATT-LCNN maintains an accuracy of 97.8% under 60dB strong noise and 97.2% at a propagation distance of 20m, while most existing SOTA methods only verify their performance under noise intensity below 50dB and propagation distance within 10m. This advantage comes from the adaptive noise suppression of the improved Kalman filter and the attention module's ability to focus on the core frequency features of leak signals that are not easily attenuated, making the algorithm more adaptable to the complex environmental conditions of municipal pipe networks.

Lightweight and Deployment Capability: The total number of parameters of ATT-LCNN is only 1.2M, with a single-sample inference time of 2.3ms on CPU, which is 86.2% smaller than the LSTM model (8.7M parameters) and 74.5% smaller than the vanilla CNN model (4.7M parameters). Most existing deep learning-based SOTA methods focus on improving accuracy while ignoring model complexity, making them difficult to deploy on embedded edge devices. In contrast, ATT-LCNN achieves a balance between high accuracy and lightweight design through depthwise separable convolution, meeting the real-time and low-computational requirements of on-site detection equipment.

5.2 Novelty and contribution of the study

The core novelty of this study lies in three aspects, which fill the gaps in existing SOTA research:

Algorithmic Novelty: An end-to-end lightweight CNN architecture fused with an attention mechanism is proposed for the first time for municipal pipeline leak acoustic signal recognition, which integrates feature extraction, feature enhancement, and classification recognition, eliminating the reliance on manual feature selection in traditional methods.

Feature Engineering Novelty: A multi-dimensional fusion feature set across time, frequency, and time-frequency domains is constructed, combined with PCA dimensionality reduction and attention enhancement, which effectively improves the representation ability of weak leak signals, providing a new solution for minor leak detection.

Engineering Practicality: The algorithm achieves a balance between detection accuracy, anti-interference performance, and lightweight design, and has been verified on simulated datasets, public field datasets, and different pipe diameters/materials, with strong generalization and robustness, which can be directly migrated to practical engineering applications.

5.3 Limitations and uncertainties of the study

Despite the excellent performance of the proposed algorithm, there are still some limitations and uncertainties that need to be addressed in future research:

Dataset Limitation: The main dataset used in this study is from a simulated pipeline device, and only supplementary verification is conducted on a public field dataset. There may be a domain gap between simulated data and real-world municipal pipe network data, which may affect the model's performance in practical applications. Real-world pipe networks have more complex conditions, such as multiple leak points, uneven pipe aging, and multi-media cross-interference, which are not fully covered in the current dataset.

Hyperparameter and Noise Bias: Although the model shows low sensitivity to core hyperparameters, the optimal hyperparameters are selected based on the simulated dataset, which may need to be adjusted when applied to different pipe network scenarios. In addition, the noise used in the experiment is simulated Gaussian noise, while the real-world environmental noise is more complex and non-stationary, which may lead to a decrease in the model's anti-interference performance.

Model Generalization Limitation: The current algorithm is only designed and verified for single-media pipe networks (water supply and gas supply), and has not been adapted to multi-media mixed pipe network scenarios. In addition, the model's performance under extremely high pressure, strong corrosion, and ultra-long propagation distance (>50m) has not been fully verified.

Overfitting Risk: Although 5-fold cross-validation and early stopping strategies are adopted to reduce the risk of overfitting, the total size of the dataset (1250 samples) is relatively modest for deep learning models, which may still lead to a certain degree of overfitting when the model is applied to scenarios that are significantly different from the training set.

6 Conclusion

This study systematically identifies acoustic signals of leaks in municipal pipeline networks and designs and verifies an original ATT-LCNN identification algorithm. The core research results are as follows: The algorithm effectively overcomes the pain points of traditional algorithms, such as the reliance on manual features, weak anti-interference ability, and imbalance between accuracy and real-time performance, through multidimensional feature fusion and attention enhancement design. It exhibits excellent performance under different working conditions, accurately capturing the weak features of small leak signals (0.1 1 L/min) with a detection rate of 98.7%. It is also adaptable to on-site scenarios with propagation distances of 1~20m and noise intensities of 20 60 dB, with a detection accuracy consistently above 97% and a repeatability error ≤0.4%. It can adapt to different pipe diameters from DN100 to DN300 and two commonly used pipe materials, steel and cast iron, achieving end-to-end accurate identification of leak signals. Experimental verification shows that compared with traditional machine learning and existing deep learning algorithms, this algorithm has significant

advantages in recognition accuracy, anti-interference ability, real-time performance, and generalization ability, and can meet the deployment requirements of on-site embedded detection equipment. This study has certain limitations, which are systematically discussed in Section 5.3. The core limitations include: first, the main experimental dataset is from a simulated pipeline device, which does not fully replicate the complex operating conditions of actual municipal pipe networks (such as multiple leak points, uneven pipe aging, and cross-interference of multiple media); second, the algorithm is only designed for single-media pipe networks, such as water and gas supply, and is not adapted to multi-media mixed pipe network scenarios; third, the lightweight design of the model still has room for improvement, and its adaptability under extremely high pressure, strong corrosion, and ultra-long propagation distance has not been fully verified; fourth, the dataset size is relatively modest, and there is a potential risk of overfitting in out-of-distribution scenarios.

Future research will focus on the above limitations: building a complex experimental platform that closely resembles actual municipal pipe networks, introducing real-world operating conditions such as multiple leak points and aging pipe materials, and collecting a large-scale field dataset to optimize the model; optimizing the algorithm architecture to expand its adaptability to multi-media mixed pipe networks; further lightweighting the model structure and combining edge computing technology to improve on-site deployment efficiency; exploring the integration of algorithms with IoT and big data technologies to construct an intelligent monitoring system for municipal pipe network leak detection, achieving early warning and accurate location of leak hazards, and promoting the improvement of the intelligent level of municipal pipe network operation and maintenance.

References

[1] Li, S., Dai, Z., Liu, M., & Cai, M. (2022). Leak identification method of water supply pipeline based on compressed sensing and least squares twin support vector machine. IEEE Sensors Journal, 23(7), 7115-7128. doi: 10.1109/JSEN.2022.3211343.

[2] Han, Y., Feng, X., & Todd, M. D. (2023). A novel methodology for quantitative identification of pipeline leakage and negative pressure wave velocity. Structural Health Monitoring, 22(4), 2267-2279. https://doi.org/10.1177/14759217221123403

[3] Ullah, N., Siddique, M. F., Ullah, S., Ahmad, Z., & Kim, J. M. (2024). Pipeline leak detection system for a smart city: leveraging acoustic emission sensing and sequential deep learning. Smart Cities, 7(4), 2318-2338. https://doi.org/10.3390/smartcities7040091

[4] Zhao, S. L., Liu, S. G., Zhao, D., Qiu, B., Hong, Z., & Shi, X. L. (2024). Experimental study on leakage location of branched pipelines based on acoustic propagation characteristics. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 238(5), 2186-2196. https://doi.org/10.1177/09544089231158953

[5] Feng, C., Zhao, J., Ran, Q., Qu, M., & Guo, Z. (2024). Acoustic-based approach for micro-leakage detection and localization in water supply pipelines. Environmental Science: Water Research & Technology, 10(8), 1881-1889. https://doi.org/10.1039/D3EW00686G

[6] Zhu, Y., Lang, X., Zhang, L., & Cai, Z. (2023). Leak localization method of jet fuel pipeline based on second-generation wavelet transform and short-time energy time delay estimation. IEEE Sensors Journal, 23(3), 2823-2832. doi: 10.1109/JSEN.2022.3233660.

[7] Kothandaraman, M., Law, Z., Ezra, M. A., Pua, C. H., & Rajasekaran, U. (2022). Water pipeline leak measurement using wavelet packet-based adaptive ICA. Water Resources Management, 36(6), 1973-1989. https://doi.org/10.1007/s11269-022-03119-y

[8] Saravanabalaji, M., Sivakumaran, N., Ranganthan, S., & Athappan, V. (2023). Acoustic signal-based water leakage detection system using hybrid machine learning model. Urban Water Journal, 20(9), 1123-1139. https://doi.org/10.1080/1573062X.2023.2239782

[9] Chen, C., Hao, P., Liu, J., Ni, L., Jiang, J., Diao, X., & Gu, W. (2023). Pipeline leak AE signal denoising based on improved SSA-K-α index-VMD-MD. IEEE Sensors Journal, 23(21), 26177-26194. doi: 10.1109/JSEN.2023.3314166.

[10] Long, Y., Zhang, J., Huang, S., Peng, L., Wang, W., Wang, S., & Zhao, W. (2022). A novel crack quantification method for ultra-high-definition magnetic flux leakage detection in pipeline inspection. IEEE Sensors Journal, 22(16), 16402-16413. doi: 10.1109/JSEN.2022.3190684.

[11] Priyanka, E. B., & Thangavel, S. (2022). Multi-type feature extraction and classification of leakage in oil pipeline network using digital twin technology. Journal of Ambient Intelligence and Humanized Computing, 13(12), 5885-5901. https://doi.org/10.1007/s12652-022-03818-9

[12] Lyu, C., Zhang, M., Li, B., Liu, Y., & Lin, X. (2022). High reliability pipeline leakage detection based on machine vision in complex industrial environment. IEEE Sensors Journal, 22(21), 20748-20760. doi: 10.1109/JSEN.2022.3206456.

[13] Abdelmageed, S., Tariq, S., Boadu, V., & Zayed, T. (2022). Criteria-based critical review of artificial intelligence applications in water-leak management.

Environmental Reviews, 30(2), 280-297. https://doi.org/10.1139/er-2021-0046

[14] Sousa, D. P., Du, R., Mairton Barros da Silva Jr, J., Cavalcante, C. C., & Fischione, C. (2023). Leakage detection in water distribution networks using machine-learning strategies. Water Supply, 23(3), 1115-1126. https://doi.org/10.2166/ws.2023.054

[15] Lu, S., Zhou, T., Wang, C., Lin, Z., & Yi, G. (2023). An internal detector positioning method in oil pipelines using vibration signal. IEEE Sensors Journal, 23(12), 13411-13421. doi: 10.1109/JSEN.2023.3273534.

[16] Huynh, C., Hibert, C., Jestin, C., Malet, J. P., Clément, P., & Lanticq, V. (2022). Real-time classification of anthropogenic seismic sources from distributed acoustic sensing data: Application for pipeline monitoring. Seismological Society of America, 93(5), 2570-2583. https://doi.org/10.1785/0220220078

[17] Goyal, N., Nain, M., Singh, A., Abualsaud, K., Alsubhi, K., Ortega-Mansilla, A., & Zorba, N. (2022). An anchor-based localization in underwater wireless sensor networks for industrial oil pipeline monitoring. IEEE Canadian Journal of Electrical and Computer Engineering, 45(4), 466-474. doi: 10.1109/ICJECE.2022.3206275.

[18] Lu, H., Xi, D., Xiang, Y., Su, Z., & Cheng, Y. F. (2025). Vehicle–canine collaboration for urban pipeline methane leak detection. Nature Cities, 2(4), 336-343. https://doi.org/10.1038/s44284-024-00183-w

[19] Abuhatira, A. A., Salim, S. M., & Vorstius, J. B. (2023). CFD-FEA based model to predict leak-points in a 90-degree pipe elbow. Engineering with Computers, 39(6), 3941-3954. https://doi.org/10.1007/s00366-023-01853-4

[20] Gemeinhardt, H., & Sharma, J. (2023). Machine-learning-assisted leak detection using distributed temperature and acoustic sensors. IEEE Sensors Journal, 24(2), 1520-1531. doi: 10.1109/JSEN.2023.3337284.

[21] Chen, M., Lu, Y., Wu, W., Ye, Y., Wei, B., & Ni, Y. (2025). Multi-Scale Frequency-Aware Transformer for Pipeline Leak Detection Using Acoustic Signals. *Sensors*, *25*(20), 6390. https://doi.org/10.3390/s25206390

[22] Fares, A., Tijani, I. A., Rui, Z., & Zayed, T. (2023). Leak detection in real water distribution networks based on acoustic emission and machine learning. *Environmental Technology*, *44*(25), 3850-3866. https://doi.org/10.1080/09593330.2022.2074320

[23] Hugar, G., & Kagalkar, R. M. (2025). Hybrid Dual-Stream Deep Learning Approach for Real-Time Kannada Sign Language Recognition in Assistive Healthcare. *J. Inf. Syst. Eng. Bus. Intell*, *11*, 393-406. http://dx.doi.org/10.20473/jisebi.11.3.393-406