ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

Gas Pipeline Leakage Detection Technology Based on Multi-model Data Fusion

Xinying Wang[1], *Yan Chen*[1,*], *Hongxin Wang*[1] *and Tingting Jiang*[1]

[1] School of Safety and Engineering, Changzhou University, Changzhou, Jiangsu, 213164, China

SUMMARY: *Quickly distinguish between a normal state and various levels of natural gas pipeline leakage in leakage detection. Single-sensor methods generally have defects in all aspects: although acoustic emission (AE) signals are highly sensitive to rapid changes in jet-flow-induced vibration, they may be easily contaminated by external noise or have a high threshold; infrared thermography can obtain direct thermal images, but it is also overly sensitive to small amounts of heat accumulation or changes in the background temperature. To overcome the above deficiencies, this paper introduces an ELM-LPP multimodal fusion model that unites AE parameters and infrared thermograms in a shared latent space. The four components of the proposed model are: an extreme learning machine (ELM), locality-preserving projection (LPP), adaptive feature weighting and joint graph regularization. The pipeline leakage experiment is a carbon-steel loop that uses air as the safe substitute medium, has three valve-opening leakage levels, an AE sensor, and an infrared thermal imager. A total of 3,485 valid AE samples and 774 infrared images were obtained, and after temporal alignment, 774 AE-infrared pairs were created for fusion learning. Evaluation of the ELM-LPP model includes confusion matrices, class-wise F1 scores, multiclass accuracy and binary-subtask accuracy. F1 scores for normal, light leakage, moderate leakage and heavy leakage were 97.03%, 97.33%, 95.16% and 97.19% for multimodal input, respectively. The Multiclass Accuracy is 96.90% and the Macro F1 score is 96.68%. Multimodal fusion reduced leakage-state misclassification and improved discrimination of leakage severity compared with AE-only and infrared-only inputs. In the binary subtasks, the proposed method achieved ACC values of 1.0000 for normal vs. light leakage, 0.9123 for light vs. moderate leakage, and 0.9357 for moderate vs. heavy leakage, and was better than AWDR, HGSCCA, MvADL and OLFG. Therefore, based on the above analysis, acoustic-infrared data fusion can offer a more stable support for the early-stage detection of pipeline leakage than either single data source.*

KEYWORDS: *Natural gas pipeline; leakage detection; acoustic emission; infrared thermography; multimodal data fusion; extreme learning machine; locality preserving projection*

1 Introduction

Natural gas has a high heat value, is relatively low in carbon emissions, and is not highly polluting; therefore, most other fossil fuels can serve as transitional fuels. Pipeline transportation is still the main way of long-distance gas supply, but with the expansion of

*fjzzyxcy@163.com
https://doi.org/10.65102/is20261271 Published: 10 June, 2026

pipeline networks, urban and industrial systems are increasingly exposed to corrosion, third-party damage, aging equipment and abnormal operating loads [1]. A leak event will cause damage to equipment and other assets, fire and explosion, and stop work in the downstream area. Therefore, the leak diagnosis should be able to determine whether a leak has occurred and identify the degree of the leak reliably enough for operational response. Studies on in-pipe isolation and emergency control have also shown that prompt identification of a leak is required for leak handling rather than detecting it after the fact [2].

At present, the general categories of current leakage detection methods are acoustic, pressure-wave, mass-balance, magnetic-flux, infrared, optical gas imaging and data-driven methods. AE sensing is now popular because when gas escapes through an opening, it produces high-frequency vibrations and acoustic signals that can be analyzed for leaks [3]. According to the research on methane emissions, in order to prevent serious consequences at a later time, any small or early-stage leak needs to be identified and addressed in time [4]. Infrared and optical methods are non-contact visual aids that can detect temperature variations or gas plumes in specific image acquisition settings [5]. Recently, AE-based deep learning research has improved leak-size classification by directly extracting features from non-stationary signals [6], and ELM-based leakage identification has demonstrated a fast learning speed in pipeline scenarios with a small computational burden [7]. AE and infrared thermography are still relatively uncertain. AE data may contain noise caused by valves, supports and compressors, and infrared images are affected by ambient temperature, emissivity and imaging distance.

Deep learning is often used for diagnostic tasks, but it requires large amounts of labelled data and has the risk of poor generalisation even when achieving high accuracy on the training set [8]. Extreme Learning Machines are very practical because the weights of the input layer and hidden layers can be set randomly and only the output weights need to be optimised analytically [9]. Related ELM studies have shown that they learn rapidly and have put forward pruning, evolutionary optimisation and other stability-enhancing strategies [10-13]. Random Fourier ELM and dynamic graph-ELM have been proposed to deal with the problems of feature redundancy and graph structure, and further development has been achieved [14, 15]. The above studies have shown that ELM is suitable for leak diagnosis, but a typical ELM fails to retain the local structure of multimodal data effectively.

Locality-preserving projection (LPP) is a linear manifold learning method that maintains the neighbourhood structure after dimension reduction [16]. Graph-regularised learning has been used to learn sample similarity and enhance robustness in cases with few labels or noisy data [17-19]. Robust discriminative classification methods have also pointed out that sample outliers and feature noise can significantly distort fixed graphs [20]. Therefore, this study does not build a single static graph for model training at present. Build a joint graph in the latent space instead and update the graph along with feature weights and ELM parameters. Recent sequential ELM studies and generalised power-iteration solvers have provided methods to address the optimisation problem under such restrictions efficiently [21, 22].

The main contribution of this paper is a leakage-diagnosis framework that integrates AE parameters and infrared thermal images via an ELM-LPP objective function. Representative multimodal fusion algorithms, such as OLFG, AWDR, HGSCCA and MvADL [23-26], are used for comparison in this study. Recently, research on intelligent pipeline monitoring and gas-leak review has focused on building multimodal, noise-tolerant and field-transferable diagnosis strategies [27, 28]. Therefore, in light of the above, this study focuses on the following three verifiable points: (1) whether AE and infrared features offer supplementary evidence for leakage severity; (2) whether latent-space graph learning enhances fusion

performance; and (3) whether the proposed method exceeds that of other multimodal classifiers on the same experimental data.

The general research path is as follows: Figure 1. The previous graphical abstract and highlight-style content have been moved from the front matter to the introduction and method description sections, and the article now begins directly with the title, abstract, and main text.

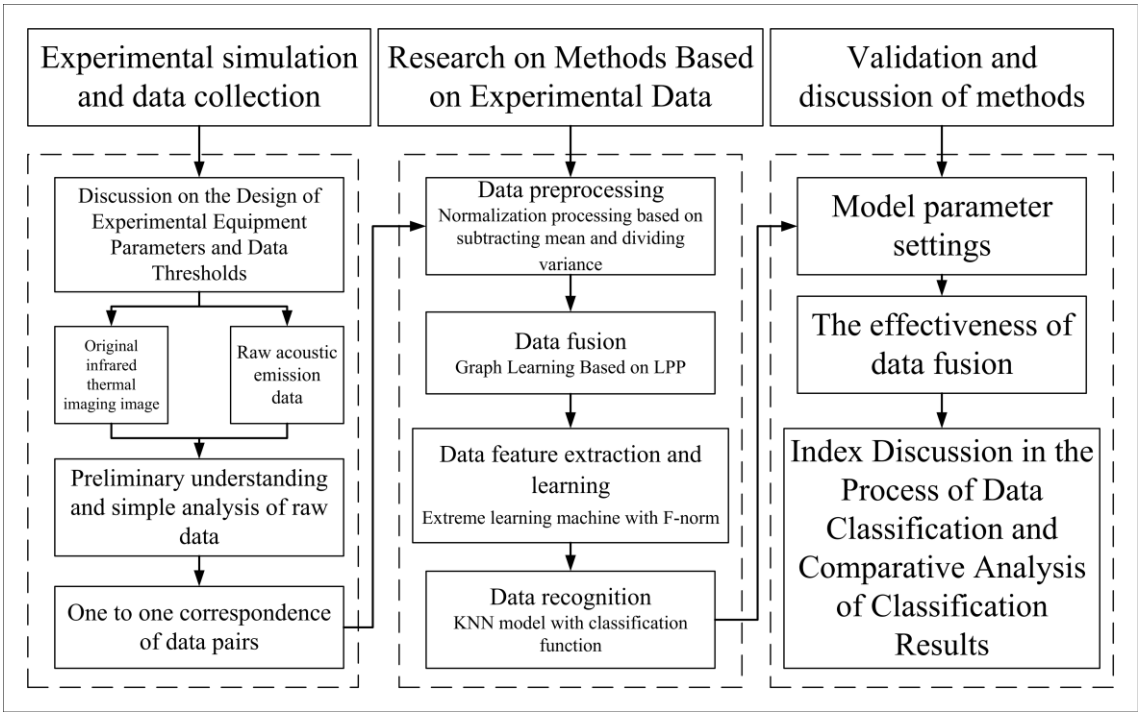

Figure 1: Revised Workflow of the Acoustic-Infrared Multimodal Leakage Diagnosis Study.

Figure 1 shows the whole path of experimental acquisition and preprocessing, latent-space fusion, and diagnostic validation. Therefore, it does not need to have an independent graphical-abstract page, and the workflow will be introduced in the main text before describing the method in detail.

2 Methods

2.1 Symbols and Definitions

Scalars, vectors, and matrices are denoted by plain lowercase letters, bold lowercase letters, and bold uppercase letters, respectively. For a matrix $\mathbf{M} \in \mathbb{R}^{d\times n}$, $\mathbf{m}_j$ and m_{ij} denote the j-th column and the (i,j)-th entry of $\mathbf{M}$, respectively. $\| \mathbf{M} \|_F$ denotes the Frobenius norm of $\mathbf{M}$. The transpose and trace of $\mathbf{M}$ are written as $\mathbf{M}^T$ and $\mathrm{Tr}(\mathbf{M})$, respectively.

A multimodal dataset is denoted as $\mathbf{X} = [\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^V]$, where V is the number of modalities. For the v-th modality, the original feature matrix is $\mathbf{X}^v = [\mathbf{x}_1^v, \mathbf{x}_2^v, \dots, \mathbf{x}_n^v] \in \mathbb{R}^{d_v\times n}$, where $\mathbf{x}_i^v$ is the i-th sample, d_v is the feature dimension, and n is the total number of samples. $\mathbf{T} \in \mathbb{R}^{c\times n}$ denotes the output matrix, with $c \le \max(d_1, d_2, \dots, d_V)$.

2.2 ELM Network and LPP Algorithm

ELM is a single-hidden-layer feedforward neural network. Set the number of hidden neurons; then, randomly initialize the hidden-layer input weights and biases, and obtain the output-

layer weights by analytical derivation. The modified ELM architecture that will be used in this paper is shown in Figure 2.

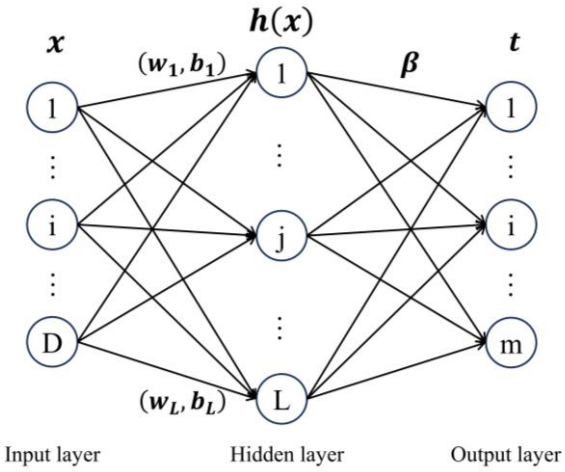

Figure 2: ELM Network Architecture for Latent Feature Mapping.

Figure 2 shows that the random hidden layer maps AE or infrared features into a nonlinear representation, and the output weight matrix beta is solved directly. This design avoids iterative training of hidden layers and is thus suitable for rapid leakage-state recognition.

Extreme Learning Machines (ELMs) are effective algorithms for single hidden-layer feedforward neural networks (SLFNs) that have shown strong generalization performance, require little manual configuration, and train quickly. After setting the number of hidden-layer neurons, ELM finds the global optimum in a single step for the hidden layer without iteration. The Design can improve the speed of learning and generalisation simultaneously. A representative SLFN generally has three parts: an input layer, a hidden layer, and an output layer, as shown in Figure 2.

The Mathematical Model of ELM is as follows:

$$t_j = \sum_{i=1}^{L} \boldsymbol{\beta}_i \, g\left(\boldsymbol{\omega}_i \mathbf{x}_j + b_i\right) \tag{1}$$

where t_j is the output corresponding to the input sample x_j; β_i is the output weight vector; $g(\cdot)$ denotes the activation function; ω_i is the input weight vector; b_i is the bias term; and L is the number of hidden-layer neurons.

Eq. (1) can be written in matrix form as:

$$\boldsymbol{\beta}^T \mathbf{H}^T = \mathbf{T} \tag{2}$$

where the hidden-layer output matrix $\mathbf{H}$ is expressed as:

$$\mathbf{H} = \begin{bmatrix} g(\boldsymbol{\omega}_1 \mathbf{x}_1 + b_1) & g(\boldsymbol{\omega}_2 \mathbf{x}_1 + b_2) & \cdots & g(\boldsymbol{\omega}_L \mathbf{x}_1 + b_L) \\ g(\boldsymbol{\omega}_1 \mathbf{x}_2 + b_1) & g(\boldsymbol{\omega}_2 \mathbf{x}_2 + b_2) & \cdots & g(\boldsymbol{\omega}_L \mathbf{x}_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g\left(\boldsymbol{\omega}_1 \mathbf{x}_q + b_1\right) & g\left(\boldsymbol{\omega}_2 \mathbf{x}_q + b_2\right) & \cdots & g\left(\boldsymbol{\omega}_L \mathbf{x}_q + b_L\right) \end{bmatrix} \tag{3}$$

Here, $\boldsymbol{\omega}_i$ and b_i are randomly initialized by the ELM algorithm. The output weight can be estimated using the Moore-Penrose generalized inverse $\mathbf{H}^+$ of the hidden-layer output matrix:

$$\widehat{\boldsymbol{\beta}} = \mathbf{H}^{+}\mathbf{T} \tag{4}$$

This solution performs well when $\mathbf{H}^T\mathbf{H}$ is nonsingular, but its generalization ability may degrade when $\mathbf{H}^T\mathbf{H}$ is singular. To improve stability, an L_2-norm regularization term is introduced into the objective function:

$$\min_{\boldsymbol{\beta}} \|\boldsymbol{\beta}^T\mathbf{H}^{vT} - \mathbf{T}\|_F^2 \tag{5}$$

Locality Preserving Projection (LPP) is a well-known way of manifold learning for dimensionality reduction and feature extraction. Reduce the size of the data to a small amount and keep the local geometric structure and neighbour relationships.

Generally speaking, LPP aims to keep nearby points in the high-dimensional space close together in a low-dimensional embedding. The following are the purpose functions:

$$\min \sum_{i,j} A_{ij} \|\mathbf{y}_i - \mathbf{y}_j\|^2 \tag{6}$$

where $\mathbf{y}_i$ and $\mathbf{y}_j$ are the low-dimensional embeddings of $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. $A_{ij} = \exp\left(-\| \mathbf{x}_i - \mathbf{x}_j \|^2/t\right)$ is the weight in the adjacency matrix, and t is the kernel width parameter. The objective encourages samples with $A_{ij} > 0$ to remain close in the latent space.

2.3 LPP-Regularised ELM Objective for Multimodal Fusion

The initial features of each modality in the multimodal pipeline leakage data are mapped to a shared latent space via ELM.

$$\min_{\mathbf{T},\boldsymbol{\beta}} \sum_{v=1}^{V} \|\boldsymbol{\beta}^T\mathbf{H}^{vT} - \mathbf{T}\|_F^2 \tag{7}$$

where $\boldsymbol{\beta} \in \mathbb{R}^{l\times c}$ is the output weight matrix, $\mathbf{H}^{vT} = g(\mathbf{W}^v\mathbf{X}^v + \mathbf{b}^v\mathbf{1}_n^T) \in \mathbb{R}^{l\times n}$ is the transposed hidden-layer output matrix of modality v, $\mathbf{W}^v \in \mathbb{R}^{l\times d_v}$ is the input weight matrix, $\mathbf{b}^v \in \mathbb{R}^l$ is the bias vector, and $\mathbf{1}_n \in \mathbb{R}^n$ is an all-ones column vector.

Eq. (7) processes the different modalities separately and does not consider cross-modal correlation. Integrate complementary information from all modalities in a shared latent space.

$$\min_{\mathbf{T},\boldsymbol{\beta},\mathbf{S},\boldsymbol{\theta}^v} \sum_{v=1}^{V} \|\boldsymbol{\beta}^T\mathbf{H}^{vT} - \mathbf{T}\|_F^2 + \alpha\|\boldsymbol{\beta}\|_F^2 \tag{8}$$

where α is a regularization parameter.

To maintain the similarity of samples in the latent space, a graph regularisation term is added:

$$\sum_{v=1}^{V}\sum_{i=1}^{n}\sum_{j=1}^{n} \left\|\mathbf{t}_i^v - \mathbf{t}_j^v\right\|_2^2 a_{ij}^v \tag{9}$$

where $a_{ij}^v = \exp\left(-\| \mathbf{x}_i^v - \mathbf{x}_j^v \|_2^2/t\right)$ denotes the similarity between samples $\mathbf{x}_i^v$ and $\mathbf{x}_j^v$ in modality v, and $\mathbf{t}_i$ is the i-th column of $\mathbf{T} = \boldsymbol{\beta}^T\mathbf{H}^{vT}$.

To solve the noise sensitivity of pre-constructed graphs and the absence of inter-modal consistency, an adaptive joint graph learning strategy is introduced:

$$\begin{aligned} &\textstyle\sum_{v=1}^{V}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\|\mathbf{t}_i^v-\mathbf{t}_j^v\right\|_2^2 a_{ij}^v+\lambda\|\mathbf{S}-\sum_{v=1}^{V}\mathbf{A}^v\|_F^2 \\ &\text{s.t.}\quad \mathbf{S1}=\mathbf{1},\quad \mathbf{S}\geq 0 \end{aligned} \tag{10}$$

where $\mathbf{S}\in\mathbb{R}^{n\times n}$ is the joint similarity matrix and λ is a regularization parameter.

Eq. (10) can be written in trace form as:

$$\begin{aligned} &\textstyle\sum_{v=1}^{V}\mathrm{Tr}\left(\boldsymbol{\beta}^T\mathbf{H}^{vT}\mathbf{L}\mathbf{H}^v\boldsymbol{\beta}\right)+\lambda\mathrm{Tr}[(\mathbf{S}-\sum_{v=1}^{V}\mathbf{A}^v)(\mathbf{S}-\sum_{v=1}^{V}\mathbf{A}^v)^T] \\ &\text{s.t.}\quad \mathbf{S1}=\mathbf{1},\quad \mathbf{S}\geq 0 \end{aligned} \tag{11}$$

where $\mathbf{L}=\mathbf{D}-\mathbf{A}$ is the Laplacian matrix of the joint graph, and $\mathbf{D}$ is a diagonal degree matrix with $D_{ii}=\sum_j a_{ij}$. The first term preserves local structure in the latent space, while the second term enforces consistency between the joint graph and individual modality graphs.

The impact of the different features on pipeline leakage diagnosis varies, and their weights also change under different leakage conditions. A adaptive feature-weighting matrix is used to obtain the feature importance:

$$\begin{aligned} &\textstyle\sum_{v=1}^{V}\mathrm{Tr}\left(\boldsymbol{\beta}^T\mathbf{H}^{vT}\mathbf{L}\mathbf{H}^v\boldsymbol{\beta}\right)+\lambda\mathrm{Tr}[(\mathbf{S}-\sum_{v=1}^{V}\mathbf{A}^v)(\mathbf{S}-\sum_{v=1}^{V}\mathbf{A}^v)^T] \\ &\text{s.t.}\quad (\boldsymbol{\theta}^v)^T\boldsymbol{\theta}^v=\mathbf{I},\quad \mathbf{S1}=\mathbf{1},\quad \mathbf{S}\geq 0 \end{aligned} \tag{12}$$

where $\boldsymbol{\theta}^v$ is a diagonal matrix whose diagonal entries θ_{ii}^v quantify the importance of the i-th feature in modality v. Larger values indicate greater contribution to latent representation and leakage diagnosis.

Combine (8) and (12) to get the final objective function as follows:

$$\begin{aligned} &\min_{\mathbf{T},\boldsymbol{\beta},\mathbf{S},\boldsymbol{\theta}^v}\textstyle\sum_{v=1}^{V}\|\boldsymbol{\beta}^T\mathbf{H}^{vT}-\mathbf{T}\|_F^2 \\ &+\gamma\textstyle\sum_{v=1}^{V}\mathrm{Tr}\left(\boldsymbol{\beta}^T\mathbf{H}^{vT}\mathbf{L}\mathbf{H}^v\boldsymbol{\beta}\right) \\ &+\lambda\mathrm{Tr}[(\mathbf{S}-\textstyle\sum_{v=1}^{V}\mathbf{A}^v)(\mathbf{S}-\sum_{v=1}^{V}\mathbf{A}^v)^T]+\alpha\|\boldsymbol{\beta}\|_F^2 \\ &\text{s.t.}\quad (\boldsymbol{\theta}^v)^T\boldsymbol{\theta}^v=\mathbf{I},\quad \mathbf{S1}=\mathbf{1},\quad \mathbf{S}\geq 0 \end{aligned} \tag{13}$$

In Eq. (13), the first term projects multimodal data into a shared latent space. The second and third terms form the adaptive joint graph regularization, allowing the joint graph $\mathbf{S}$ to be learned during optimization. The last term regularizes the output weights to reduce overfitting. The learned graph is more robust to noise and better suited for leakage diagnosis.

The optimised path of Eq. (13) is shown in Figure 3. The figure has been added to the main text, and its caption has been modified to place the introduction and explanation of the objective-flow figure before and after its display, respectively.

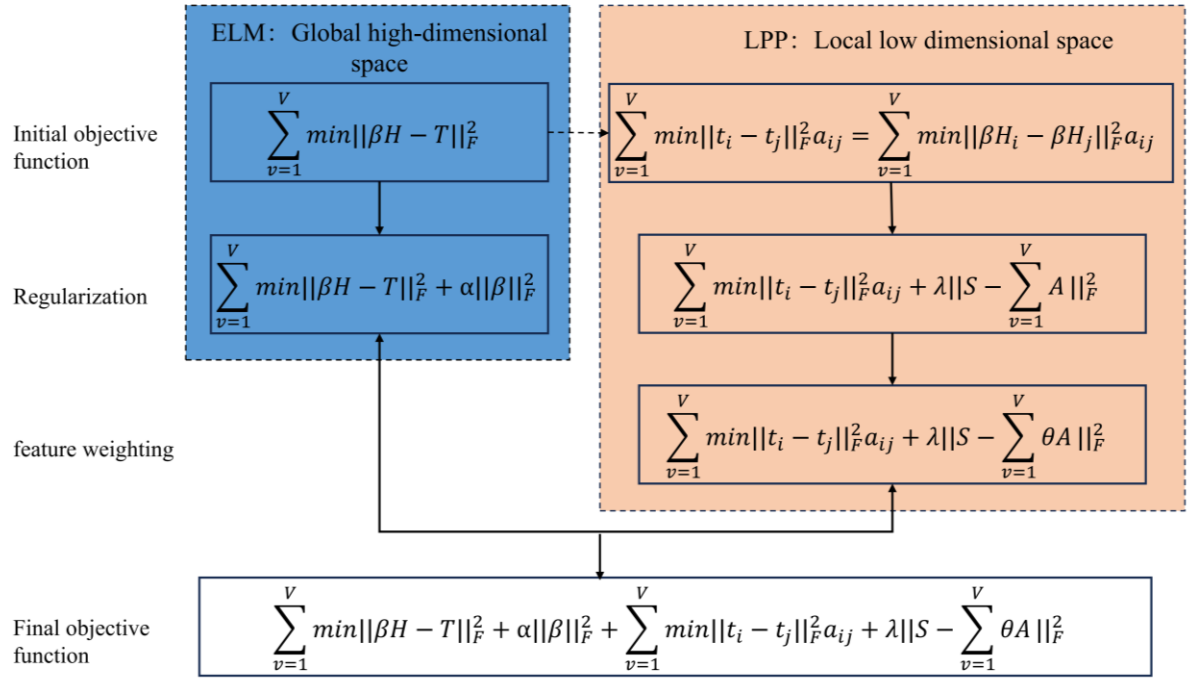

Figure 3: Objective-function Construction for the ELM-LPP Fusion Model.

Figure 3 shows that the final objective is obtained by adding regularization, LPP-based local preservation, adaptive graph learning and feature weighting to the basic ELM reconstruction term. Thus, this model can handle the high dimensionality and inhomogeneity of AE-infrared data directly.

2.4 Optimisation Algorithm

Eq. (13) is optimized using an alternating optimization strategy combined with the generalized power iteration (GPI) algorithm. The variables $\boldsymbol{\beta}$, $\mathbf{T}$, $\boldsymbol{\theta}^{v}$, and $\mathbf{S}$ are updated alternately while the remaining variables are fixed.

With $\mathbf{T}$, $\boldsymbol{\theta}^{v}$, and $\mathbf{S}$ fixed, Eq. (13) reduces to:

$$\min_{\boldsymbol{\beta}} \sum_{v=1}^{V} \|\boldsymbol{\beta}^{T}\mathbf{H}^{vT} - \mathbf{T}\|_{F}^{2} + \gamma \sum_{v=1}^{V} \mathrm{Tr}\,(\boldsymbol{\beta}^{T}\mathbf{H}^{vT}\mathbf{L}\mathbf{H}^{v}\boldsymbol{\beta}) + \alpha\|\boldsymbol{\beta}\|_{F}^{2} \tag{14}$$

which is the same as:

$$\min_{\boldsymbol{\beta}} \mathrm{Tr}(\boldsymbol{\beta}^{T}\mathbf{N}\boldsymbol{\beta}) - 2\mathrm{Tr}(\mathbf{B}\boldsymbol{\beta}) \tag{15}$$

where $\mathrm{N} = \sum_{\mathrm{v}=1}^{\mathrm{V}}(\mathrm{H}^{\mathrm{vT}}\mathrm{H}^{\mathrm{v}} + \gamma\mathrm{H}^{\mathrm{vT}}\mathrm{L}\mathrm{H}^{\mathrm{v}}) + \alpha\mathrm{I}$ and $\mathrm{B} = \sum_{\mathrm{v}=1}^{\mathrm{V}} \mathrm{T}\,\mathrm{H}^{\mathrm{v}}$.

Differentiating and setting the result to zero yields:

$$2\mathbf{N}\boldsymbol{\beta} - 2\mathbf{B}^{T} = 0 \tag{16}$$

Thus, the optimal $\boldsymbol{\beta}$ is obtained as:

$$\boldsymbol{\beta} = \mathbf{N}^{-1}\mathbf{B}^{T} \tag{17}$$

Keeping the other variables fixed, Eq. (13) is:

$$\min_{\mathbf{T}} \sum_{v=1}^{V} \|\boldsymbol{\beta}^T \mathbf{H}^{vT} - \mathbf{T}\|_F^2 \tag{18}$$

which can be written as

$$\min_{\mathbf{T}} \mathrm{Tr}(\mathbf{TC}) + \mathrm{Tr}(\mathbf{TT}^T) \tag{19}$$

where $\mathrm{C} = -\sum_{\mathrm{v}=1}^{\mathrm{V}} 2\,\mathrm{H}^{\mathrm{v}}\beta$.

Deriving the optimal solution results in:

$$\mathbf{T} = -\frac{1}{2}\mathbf{C}^T \tag{20}$$

with $\boldsymbol{\beta}$, $\mathbf{T}$, and $\mathbf{S}$ fixed, the optimization of $\boldsymbol{\theta}^v$ becomes:

$$\begin{aligned} \min_{\boldsymbol{\theta}^v} \quad & \lambda \mathrm{Tr}[(\mathbf{S} - \textstyle\sum_{v=1}^{V} (\boldsymbol{\theta}^v)^T \mathbf{A}^v)(\mathbf{S} - \textstyle\sum_{v=1}^{V} (\boldsymbol{\theta}^v)^T \mathbf{A}^v)^T] \\ \text{s.t.} \quad & (\boldsymbol{\theta}^v)^T \boldsymbol{\theta}^v = \mathbf{I} \end{aligned} \tag{21}$$

Simplified form:

$$\begin{aligned} \min_{\boldsymbol{\theta}^v} \quad & \lambda \mathrm{Tr}[(\boldsymbol{\theta}^v)^T \mathbf{Q}^v \boldsymbol{\theta}^v - 2(\boldsymbol{\theta}^v)^T \mathbf{F}^v] \\ \text{s.t.} \quad & (\boldsymbol{\theta}^v)^T \boldsymbol{\theta}^v = \mathbf{I} \end{aligned} \tag{22}$$

where $\mathrm{Q}^{\mathrm{v}} = \mathrm{A}^{\mathrm{v}}(\mathrm{A}^{\mathrm{v}})^{\mathrm{T}}$, $\mathrm{F}^{\mathrm{v}} = \mathrm{A}^{\mathrm{v}}\mathrm{K}^{\mathrm{v}}$, and $\mathrm{K}^{\mathrm{v}} = \mathrm{S} - \sum_{\mathrm{k}\neq\mathrm{v}} \left(\theta^{\mathrm{k}}\right)^{\mathrm{T}} \mathrm{A}^{\mathrm{k}}$. This problem conforms to the GPI framework.

With β, T, and θ^{v} fixed, and letting $\mathrm{T} = \beta^{\mathrm{T}}\mathrm{H}^{\mathrm{v}\mathrm{T}}$ and $\mathrm{M} = \sum_{\mathrm{v}=1}^{\mathrm{V}} (\theta^{\mathrm{v}})^{\mathrm{T}}\,\mathrm{A}^{\mathrm{v}}$, Eq. (13) becomes:

$$\begin{aligned} \min_{\mathbf{S}} \quad & \frac{\gamma}{\lambda} \textstyle\sum_{v=1}^{V} \mathrm{Tr}\,(\mathbf{TLT}^T) + \|\mathbf{S} - \mathbf{M}\|_F^2 \\ \text{s.t.} \quad & \mathbf{S1} = \mathbf{1}, \quad \mathbf{S} \geq 0 \end{aligned} \tag{23}$$

which can be written as:

$$\min_{\mathbf{s}_i^T \mathbf{1}=1,\, \mathbf{s}_i \geq 0} \sum_{i,j} \left(s_{ij} - m_{ij}\right)^2 + \frac{\gamma}{2\lambda} \sum_{i,j} s_{ij} \sum_{v=1}^{V} \left\|\mathbf{t}_i^v - \mathbf{t}_j^v\right\|_2^2 \tag{24}$$

Since the problem is separable over i, each $\mathbf{s}_i$ can be solved independently:

$$\begin{aligned} \min_{\mathbf{s}_i} \quad & \left\|\mathbf{s}_i - \left(\mathbf{m}_i - \frac{\gamma}{4\lambda}\mathbf{e}_i\right)\right\|_2^2 \\ \text{s.t.} \quad & \mathbf{s}_i^T \mathbf{1} = 1, \quad \mathbf{s}_i \geq 0 \end{aligned} \tag{25}$$

where $e_{ij} = \sum_{v=1}^{V} \left\|\mathbf{t}_i^v - \mathbf{t}_j^v\right\|_2^2$.

The Lagrangian Function is:

$$\mathcal{L}(\mathbf{s}_i, \eta, \boldsymbol{\varepsilon}_i) = \frac{1}{2}\left\|\mathbf{s}_i - \left(\mathbf{m}_i - \frac{\gamma}{4\lambda}\mathbf{e}_i\right)\right\|_2^2 + \eta(\mathbf{s}_i^T \mathbf{1} - 1) - \boldsymbol{\varepsilon}_i^T \mathbf{s}_i \tag{26}$$

The Optimal Solution will be obtained when the Karush-Kuhn-Tucker conditions hold.

$$\mathbf{s}_i = \max\left(\mathbf{m}_i - \frac{\gamma}{4\lambda}\mathbf{e}_i + \eta\mathbf{1}, \mathbf{0}\right) \tag{27}$$

Algorithm 1. Generalized Power Iteration (GPI)

Input: symmetric matrix $\mathbf{D}^v$ and matrix $\mathbf{F}^v$. Initialize $\boldsymbol{\theta}^v$ randomly. Initialize λ so that $\mathbf{O} = \lambda\mathbf{I}_{d_v} - \mathbf{D}^v \succ 0$. While the algorithm is not converged, update $\mathbf{P} = 2\mathbf{O}\boldsymbol{\theta}^v + 2\mathbf{F}^v$. Compute the compact singular value decomposition $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \mathbf{P}$, and then update $\boldsymbol{\theta}^v = \mathbf{U}\mathbf{V}^T$. Return $\boldsymbol{\theta}^v$.

Algorithm 2. Alternating Optimisation Algorithm

Input: multimodal pipeline leakage data $\mathbf{X}^v = [\mathbf{x}_1^v, \ldots, \mathbf{x}_n^v]$, similarity matrices $\mathbf{A}^v$, and parameters α, γ, and λ. Randomly initialize $\mathbf{W}^v$ and $\mathbf{b}^v$ to compute $\mathbf{H}^v$. While the algorithm is not converged, update $\boldsymbol{\beta}$ using Eq. (17), update $\mathbf{T}$ using Eq. (20), update $\boldsymbol{\theta}^v$ using Algorithm 1 for each modality, and update $\mathbf{s}_i$ using Eq. (27) for each sample. The final output is $\boldsymbol{\beta}$.

3 Experiments and Data Preparation

3.1 Pipeline Leakage Experimental Platform

A leak experiment was conducted on a No.20 carbon steel pipe with seamless steel pipes. The three-layer spiral coiled structure of the system has a layer spacing of 320 mm. The pipeline is about 50 metres long, 2 metres wide and 1 metre high; the diameter of the pipe is 3 cm and the wall thickness is 5 mm. Compressed air was used as the experimental medium out of safety concerns, and natural gas was not. The laboratory pipeline platform is as follows: Figure 4.

Figure 4: Pipeline Leakage Experimental Simulation Platform.

Figure 4 is the coiled pipeline, air-supply branch and test area. A controlled platform can be used to keep the leakage level, sensor distance and acquisition interval the same in several trials.

A MISTRAS R31-AST sensor was selected as the AE sensor, set to an operating frequency of 10-40kHz, and used in a temperature range of -35 ° C to 75 ° C. FLUKE Ti400

thermal imager will be used for infrared imaging. The settings of the instrument used in the experiment are as follows: Table 1.

Table 1: Instrument Setting Parameters.

Parameter	Value
Acoustic-emission threshold	45 dB
Infrared thermal-imager capture interval	3 s
Distance between AE sensor and leakage source	20 cm, 60 cm, 100 cm, 120 cm
Distance between thermal imager and leakage source	75 cm

Table 1 shows that the AE threshold is used to suppress background noise, and the distance of the thermal imager is fixed to achieve stable focusing. Different distances from the AE sensor are selected to study the effect of a shift in the source or sensor position on signal reception.

Increase the Valve Opening to increase the Leakage Severity. The four operating states were: normal operation at 0°, light leakage at 0°-15°, moderate leakage at 15°-45°, and heavy leakage at 45°-60°. The valve opening conditions are as follows: Figure 5.

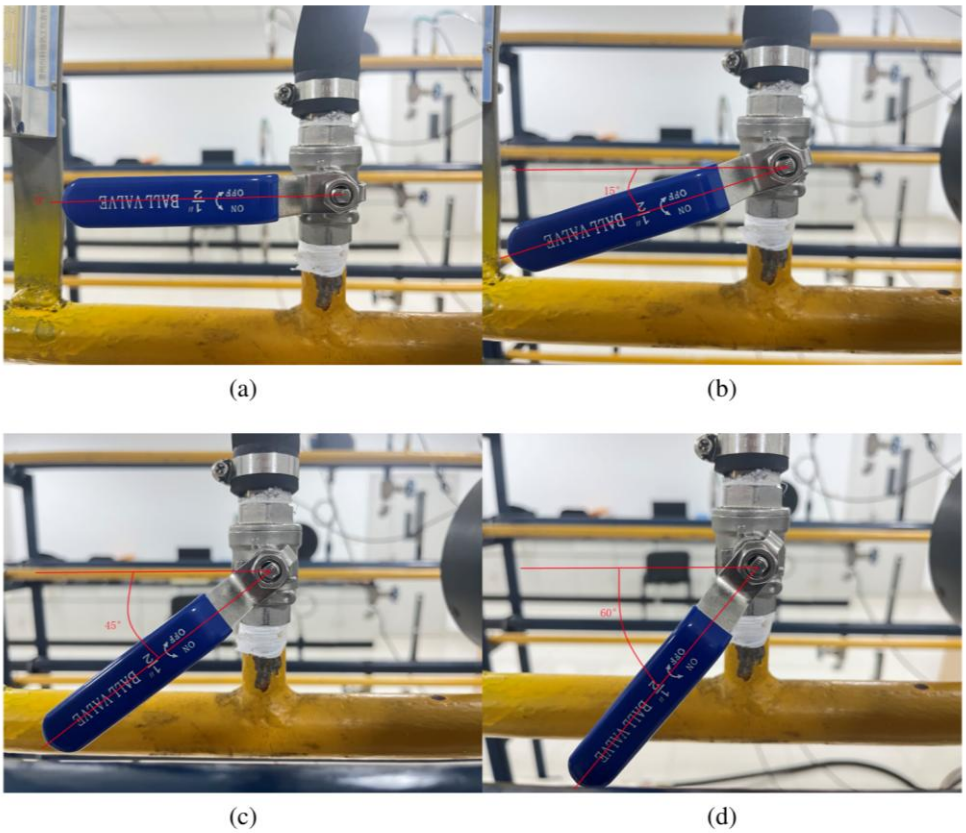

Figure 5: Valve-opening settings for leakage-level simulation: (a) normal, 0° (b) light leakage, 0°-15° (c) moderate leakage, 15°-45° (d) heavy leakage, 45°-60°.

Figure 5 presents the physical foundation of the class labels used in the classification experiment below. As the opening angle directly affects the size of the leakage aperture, it can be used to obtain light, medium and heavy leakage samples consistently.

Compressor-pressurise the pipeline to 0.6 MPa in each trial and then shut off. The valve opening was set to the target position, and at the same time, the AE sensor and infrared thermal imager collected data until the internal pressure reached atmospheric pressure. Three times in a row, the same leakage condition was repeated to reduce the impact of experimental error.

3.2 Preliminary Analysis of AE and Infrared Data

A total of 3,485 AE samples and 774 infrared thermal images were collected in the experiment. The AE parameters are rise time, count, energy, duration, amplitude, average frequency, RMS, ASL, peak frequency, inverse frequency, initial frequency, signal strength, absolute energy and frequency. To prevent the loss of information due to manual feature screening, all parameters were kept for the following fusion. The normalized means of the AE parameters under different leakage conditions are shown in Figure 6.

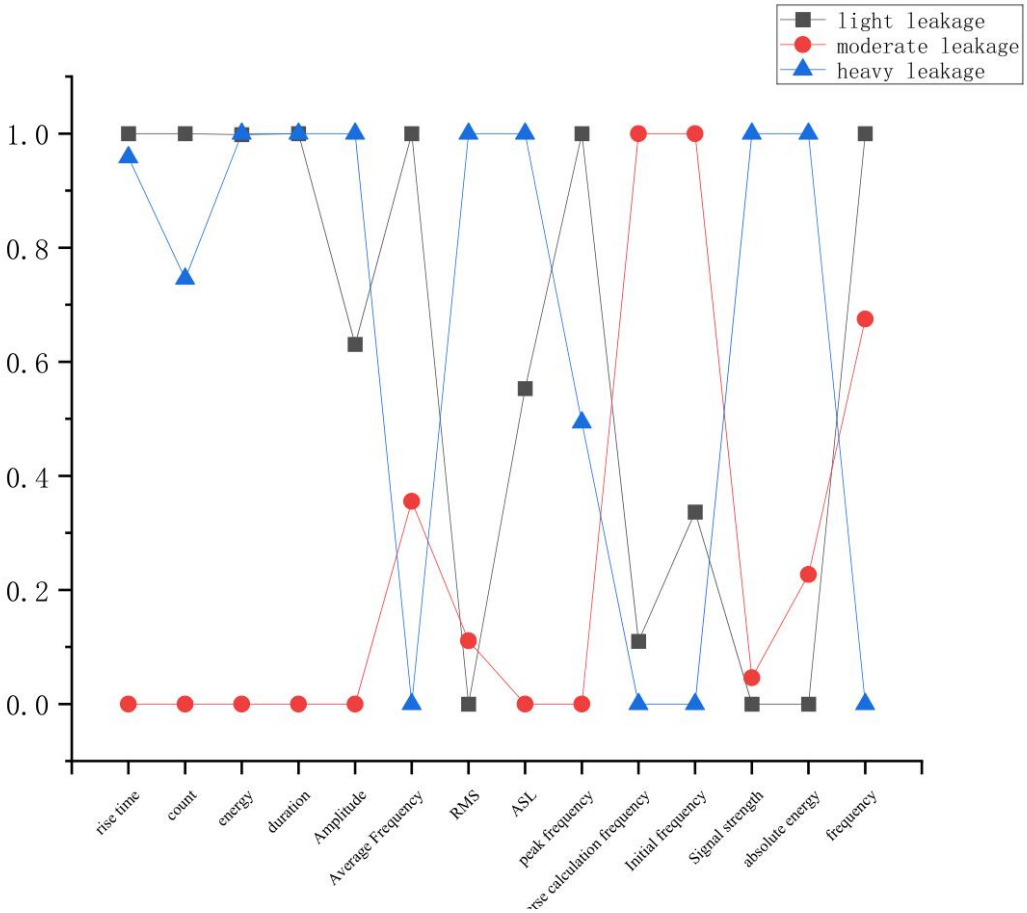

Figure 6: Comparison of Normalised Mean AE Parameters Under Different Leakage Conditions.

As shown in Figure 6, none of the AE parameters can distinguish all leakage levels monotonically. Rise time and peak frequency are high in the case of a heavy leakage; inverse and initial frequencies are relatively more pronounced under moderate leakage; and amplitude-related parameters are more sensitive to light leakage. Thus, a few features are no longer required to perform multi-dimensional feature learning.

Perform Pearson correlation analysis on the correlations of the AE parameters. The Correlation matrix is as follows: Figure 7.

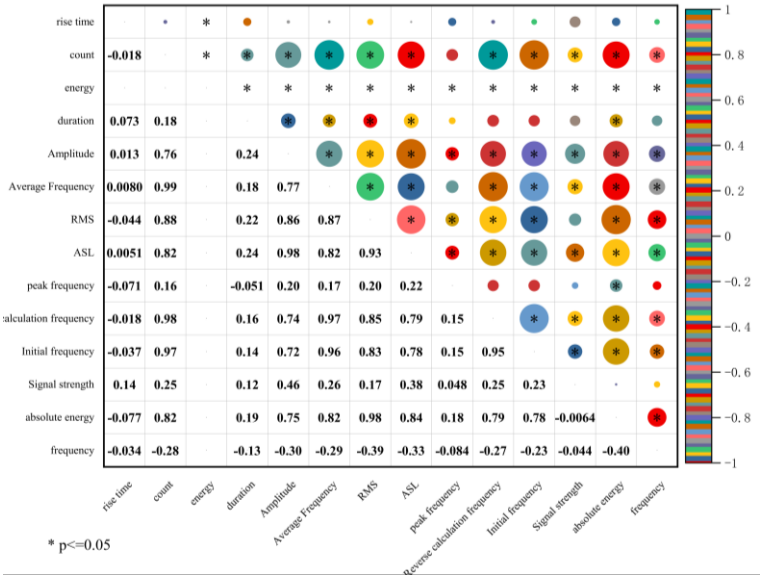

Figure 7: Pearson Correlation Analysis of AE Parameters.

As shown in Figure 7, the amplitude-related variables of energy, RMS, ASL and absolute energy show strong positive correlations among themselves, and peak frequency is negatively correlated with all these amplitude variables. The time-related indicators of rise time and duration are less correlated with the other groups. Based on the above results, it can be seen that the parameter groups are redundant and the other two are complementary.

Analysis of the infrared thermography shows both the development stage and the severity of the leakage. The first few representative thermal images are shown in Figure 8.

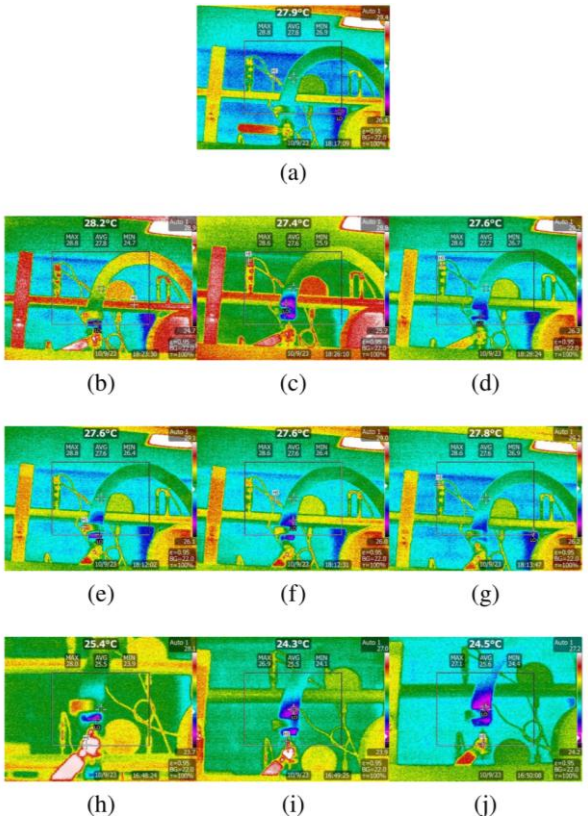

Figure 8: Infrared Thermograms at Different Stages and Leakage Severities.

Figure 8 shows that light leakage results in a small thermal spot and weak colour-scale variation; moderate leakage causes a rapid local rise in temperature and expansion of the thermal spot; and heavy leakage forms an extended area of high temperature. Infrared thermography thus provides spatial support for the time-frequency information in AE signals.

3.3 Data preprocessing and Alignment

Z-score normalisation is used to reduce the dimension problem of AE parameters and infrared-image vectors:

$$z = \frac{X-\mu}{\sigma} \tag{28}$$

where X is the original value, mu is the sample mean, and sigma is the standard deviation. Each infrared image was resized to 16 x 16 and then converted into a 1 x 256 vector so that the two-dimensional thermal image could be aligned with one-dimensional AE features.

Since the number of infrared images was less than that of the AE samples, the infrared frames were taken as the time reference. The nearest synchronous AE sample was paired with each infrared frame, and 774 multimodal samples were obtained. The Alignment process is as follows: Figure 9.

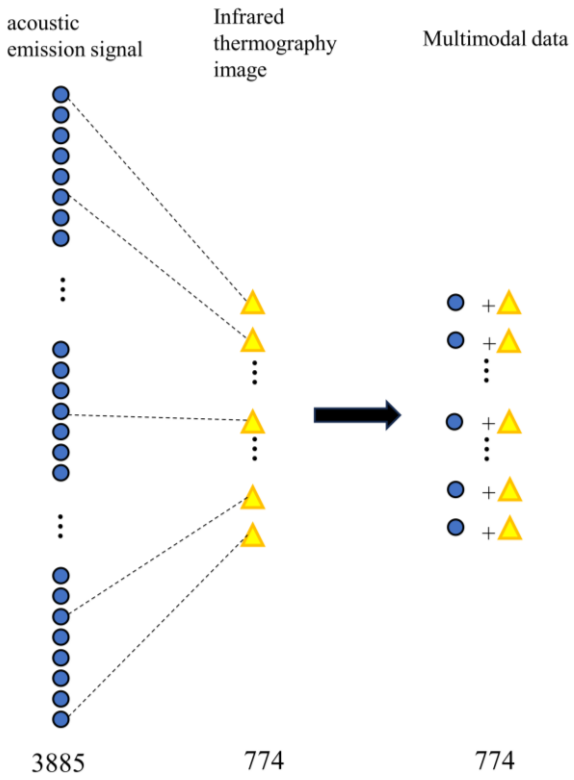

Figure 9: Timestamp-based Alignment of AE Samples and Infrared Thermal Images.

As shown in Figure 9, the fusion dataset is not a random resampled set. Each infrared frame is a single AE feature vector at the same leakage stage, and thus maintains chronological consistency between the two modalities.

The AE data for single-modal model training were split into training and validation sets at an 8:2 ratio. To balance the normal class, 400 normal-state AE samples of background operation were added to provide 3,108 AE training samples and 777 AE validation samples.

Divide the infrared data into a 5:5 ratio for training and validation, and obtain 387 training and 387 validation images. The multimodal paired data were also split at a 5:5 ratio to get 387 training pairs and 387 validation pairs.

3.4 Comparison Algorithms and Evaluation Indicators

The proposed ELM-LPP model was compared with AWDR, HGSCCA, MvADL and OLFG. The above methods were selected because they belong to different categories of multimodal fusion: adaptive weighting and discriminative regression, sparse canonical correlation with hypergraph modelling, dictionary learning and multi-view analysis, and orthogonal latent-space graph learning. The same training-validation divisions were used for all the methods.

Accuracy (ACC), sensitivity (SEN), specificity (SPE), area under the ROC curve (AUC), F1 score and macro-F1 were used to assess the performance of classification. Binary leakage detection treats leakage as the positive class and normal conditions as the negative class:

$$\mathrm{ACC} = \frac{TP+TN}{TP+TN+FP+FN} \tag{29}$$

$$\mathrm{SEN} = \frac{TP}{TP+FN} \tag{30}$$

$$\mathrm{SPE} = \frac{TN}{TN+FP} \tag{31}$$

$$F1 = \frac{2TP}{2TP+FP+FN} \tag{32}$$

$$\text{Macro-}F1 = \frac{1}{n}\sum_{i=1}^{n} F\,1_i \tag{33}$$

The original formula label for total accuracy has been replaced by macro-F1 because Equation (33) now computes the average F1 scores across all classes rather than counting all correct predictions. The correction will align the results with those in the confusion matrix in Section 4.

4 Results and Discussion

4.1 Parameter Settings

The manually selected ELM structural parameter of the ELM-LPP model is the number of hidden-layer neurons d, and the three regularization parameters are alpha, gamma, and lambda. The six candidate values for d were 10, 50, 100, 200, 500 and 1000. The diagnostic indicators for the various d values are shown in Figure 10.

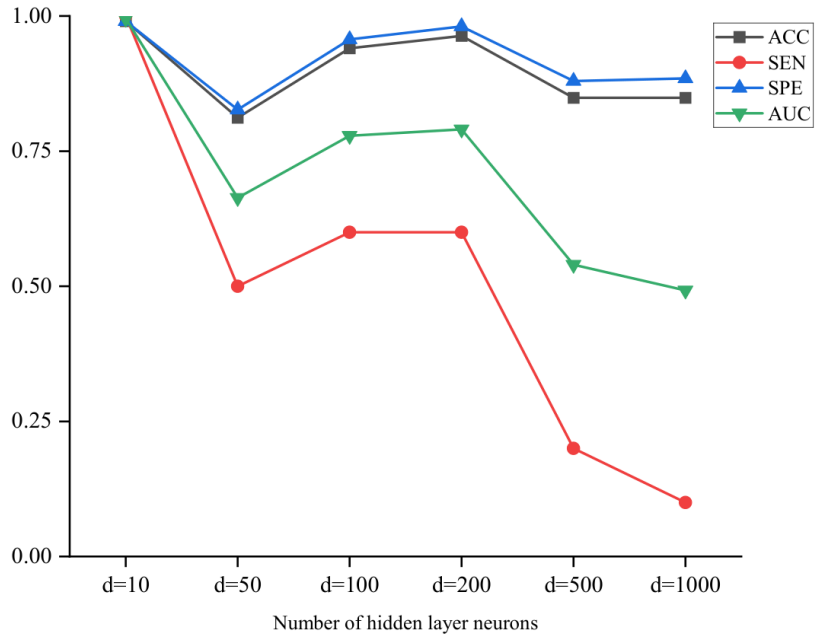

Figure 10: Influence of the number of hidden-layer neurons on Diagnostic Performance.

Figure 10 shows that the model has the best general performance at d = 10. Larger hidden layers do not improve performance and may reduce sensitivity; therefore, an excess of hidden neurons has been added as redundant projections at the present scale of data. Thus, d=10 was selected for the following experiments.

Fixing d, the regularization parameters α, γ, and λ were searched in the logarithmic interval $\{10-3,10-2,\ldots,10^{-2}\}$. The effect of the parameters is shown in Figure 11.

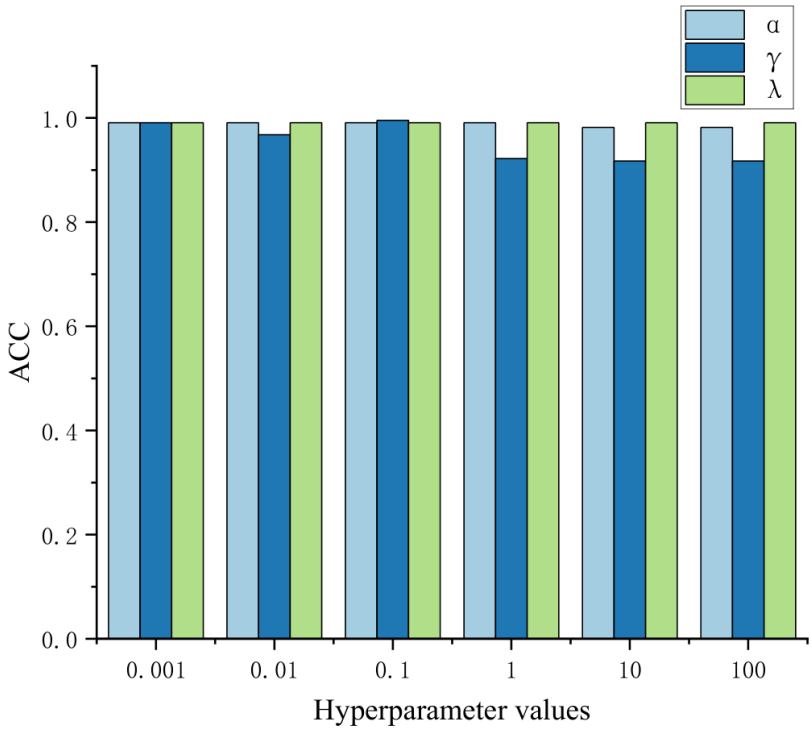

Figure 11: Influence of Alpha, Gamma and Lambda on Model Accuracy.

As shown in Figure 11, the model is relatively more sensitive to gamma than to lambda. Accuracy is still stable with all three parameters less than 1; an increase in alpha and gamma will progressively reduce the result. Based on the above trend, alpha = 0.1, gamma = 0.1, and lambda = 0.1 were finally selected for the model.

4.2 Leakage Diagnosis Process

Based on the above parameter settings, the whole ELM-LPP diagnosis process is shown in Figure 12.

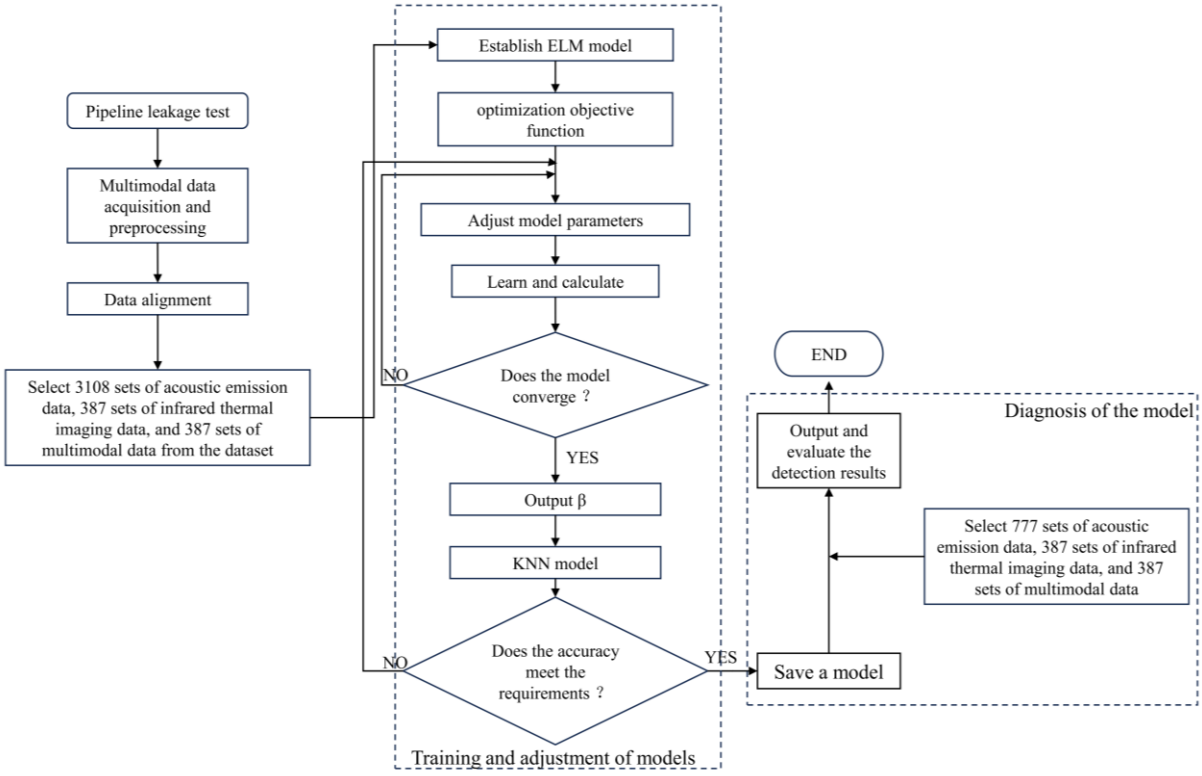

Figure 12. ELM-LPP Pipeline Leakage Diagnosis Process.

Figure 12 shows the training and validation stages. Multimodal samples are used in training to learn beta and the joint graph. Validation uses the learned representation in a KNN classifier, and the resulting output labels are compared with the four known leakage states.

4.3 Comparison of Confusion Matrices for Unimodal and Multimodal Inputs

The classification results of AE-only, infrared-only and AE-infrared inputs are shown in the confusion matrices in Figure 13. The four labels are normal operation, light leakage, moderate leakage and heavy leakage.

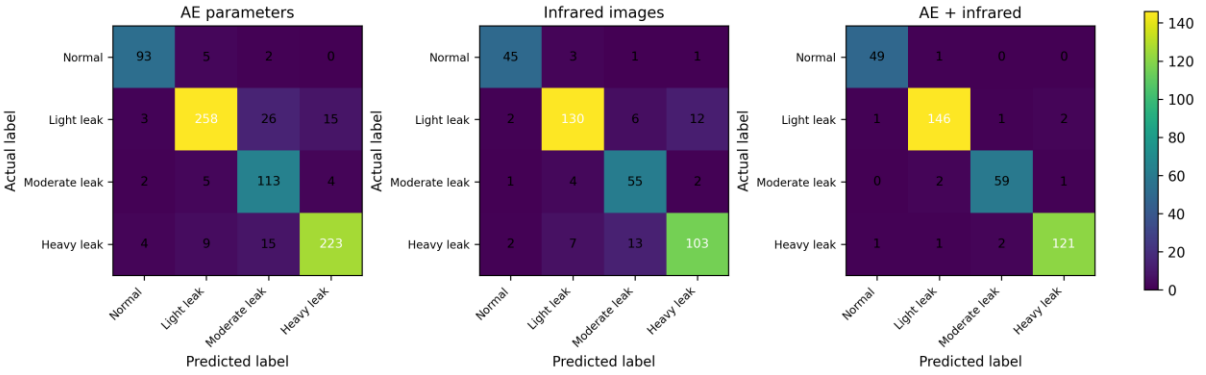

Figure 13: Confusion Matrices for AE-only, Infrared-only, and Multimodal AE-Infrared Classifications.

Figure 13 shows that AE parameters can identify the normal state well, but light, moderate and heavy leakage are confused. Infrared images reduce some of the confusion at the severity level but still fail to distinguish heavy leakage from moderate leakage. The two types of errors are both reduced: only one normal sample is misclassified as light leakage, and the three leakage severity cases have fewer cross-class errors than either unimodal input.

Confusion matrices can also serve a practical purpose for fusion. Binary classification of the normal and leakage states has a misclassification rate of 2.06% for AE input, 2.58% for infrared input, and 0.78% for multimodal input. Multimodal fusion for severity classification achieved a 98.00% correct classification rate for light leakage, a 95.16% F1 score for moderate leakage, and a 97.19% F1 score for heavy leakage. The above improvements are consistent with the complementary sensing mechanisms shown in Figures 6-8.

4.4 F1-score and Generalization Performance Analysis

Class-balanced comparison F1 scores based on confusion matrices were computed. The results are as follows: Figure 14.

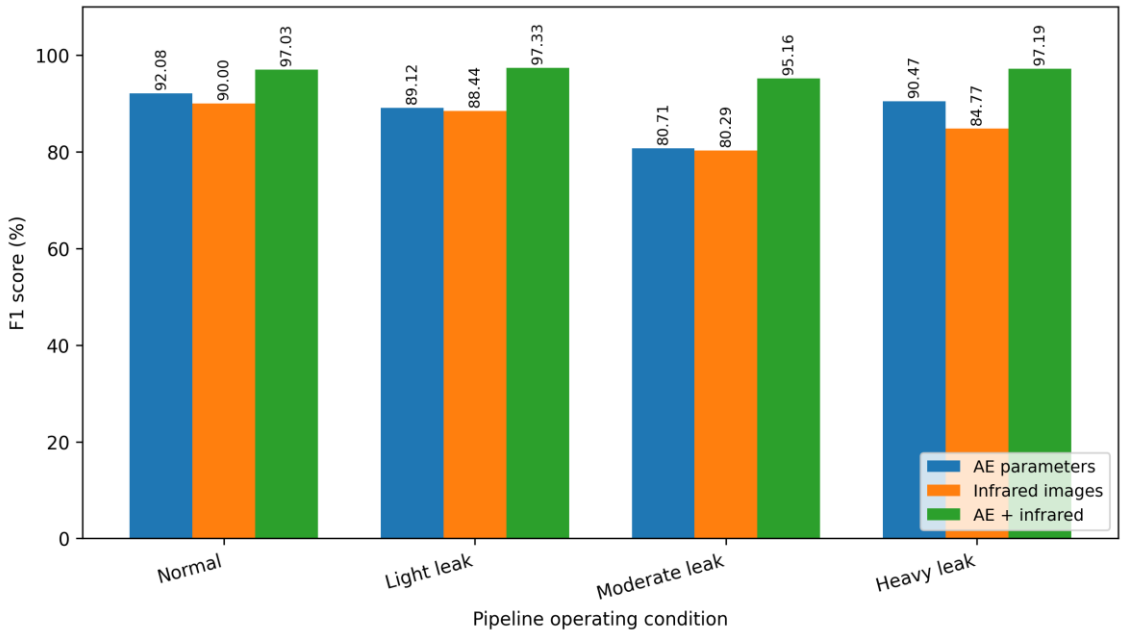

Figure 14: F1 Scores of Classes Based on Different Input Modalities.

As shown in Figure 14, multimodal input has achieved the highest F1 score for all four operating modes. AE-only input achieves the following F1 scores: 92.08%, 89.12%, 80.71%, and 90.47%, with a macro-F1 of 88.09%. The infrared-only input obtained 90.00%, 88.44%, 80.29% and 84.77%, and had a macro-F1 of 85.88%. AE-infrared input has obtained 97.03%, 97.33%, 95.16%, and 97.19%, with a macro-F1 score of 96.68% and a multi-class accuracy of 96.90%.

The highest increase in the range of moderate leakage is as follows: the F1 score of both unimodal and multimodal inputs is around 80%, and that of multimodal input reaches as high as 95.16%. Therefore, the class of moderate leakage is a difficult transition case that is more directly supported by both AE feature variation and thermal-spot evolution. It can be seen that multimodal fusion reduces false alarms of emergencies and failure to recognise increased severity during diagnosis in practice.

4.5 Comparison with Multimodal Fusion Baselines

Ten-fold cross-validation was employed to compare ELM-LPP with AWDR, HGSCCA, MvADL and OLFG on three binary classification tasks. Table 2 shows the mean ACC values.

Table 2: Comparison of ACC Values for Binary Leakage-Severity Subtasks.

Method	Normal vs. light	Light vs. moderate	Moderate vs. heavy
AWDR	1.0000	0.7807	0.9286
HGSCCA	1.0000	0.6140	0.9143
MvADL	0.6221	0.8728	0.8500
OLFG	0.9954	0.8421	0.8357
ELM-LPP	1.0000	0.9123	0.9357

As shown in Table 2, ELM-LPP has achieved the highest ACC of the three binary sub-tasks. The ordinary and light tasks are relatively easy because the background of the normal-state AE is significantly different from the leakage signal. Light-to-moderate and moderate-to-heavy tasks are more difficult because the leakage features change continuously with an increase in valve opening. ELM-LPP performs better in the above transition tasks because the adaptive joint graph maintains local class boundaries in the latent space rather than directly concatenating raw features.

ELM-LPP does not need to select many features or learn feature weights during latent-space construction, as SVM-based or dictionary-learning baselines do. This will reduce the risk of excluding useful cross-modal information in the classification. The learned joint graph is more suitable for heterogeneous AE-infrared samples than fixed-graph methods, and it updates the similarity relationship with the latent representation. The above reasons are the cause of the continuous rise in Table 2.

4.6 Limitations and Future Work

Several Deficiencies Remain. First of all, the experiment used compressed air instead of methane or natural gas for safety; therefore, the field-scale thermal and acoustic responses may differ from those in the laboratory. Second, the normal-state AE class includes background samples added for class balancing; future work should increase the number of measured no-leak records rather than using a small baseline augmentation. Third, some of the photo-based figures in the current revision have been replaced by high-resolution crops or redrawn process diagrams, but the original high-resolution camera images should still be provided before journal submission to prevent raster degradation during production. Finally,

the current model was run on a single pipeline platform. At various pipe sizes and pressures, for different shapes of leak holes, and in all environments, cross-platform verification must be carried out.

5 Summary

An ELM-LPP multimodal data-fusion method for natural gas pipeline leakage detection based on AE parameters and infrared thermal images is proposed in this paper. The graphical abstract and highlights have been moved out of the front matter and added to the main text. The modified way combines latent-space learning, adaptive joint graph construction, feature weighting and KNN classification explicitly.

The three results are as follows: First, AE and infrared thermography offer supplementary data. AE parameters can be used to identify deviations from the normal operating conditions, and infrared images can be employed to locate areas in the image with an abnormal colour or size of thermal spots. Second, Multimodal Fusion Improves Severity Classification. Based on the validation confusion matrix, the F1 scores of the multimodal model for normal, light leakage, moderate leakage and heavy leakage are 97.03%, 97.33%, 95.16 and 97.19%, respectively. The multi-class accuracy is 96.90% and the macro-F1 is 96.68%. Third, the proposed ELM-LPP model is superior to AWDR, HGSCCA, MvADL and OLFG in the binary leakage-severity subtasks, especially in the light-to-moderate and moderate-to-severe transition tasks.

In short, the results show that AE-infrared fusion can reduce the number of false leakage-level judgments compared to either modality independently. Therefore, the proposed framework is a feasible basis for laboratory leakage diagnosis and offers technical references for the development of multimodal pipeline monitoring systems. Future work should expand the dataset with field natural-gas experiments, replace all photo-based figures with original high-resolution images, and verify model robustness under different pressure and temperature conditions of the pipeline structure.

Data Availability Statement

The data, models and code for the results of this study are available from the corresponding author upon reasonable request.

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or other close personal and professional relationships that could have affected the research in this paper.

References

[1] Wang, J., Ren, L., Jia, Z., Jiang, T., & Wang, G. (2022). A novel pipeline leak detection and localization method based on the FBG pipe-fixture sensor array and compressed sensing theory. Mechanical Systems and Signal Processing, 169, 108669.

[2] Miao, X., & Zhao, H. (2022). Regulating control of in-pipe intelligent isolation plugging tool based on adaptive dynamic programming. Journal of Pipeline Systems

Engineering and Practice, 13(2), 04022003.

[3] Ullah, N., Ahmed, Z., & Kim, J.-M. (2023). Pipeline leakage detection using acoustic emission and machine learning algorithms. Sensors, 23(6), 3226.

[4] Alvarez, R. A., Pacala, S. W., Winebrake, J. J., Chameides, W. L., & Hamburg, S. P. (2012). Greater focus needed on methane leakage from natural gas infrastructure. Proceedings of the National Academy of Sciences, 109(17), 6435-6440.

[5] Naranjo, E., & Baliga, S. (2012). Early detection of combustible gas leaks using open path infrared gas detectors. In Advanced Environmental, Chemical, and Biological Sensing Technologies IX (Vol. 8366, pp. 237-242). SPIE.

[6] Saleem, F., Ahmad, Z., Siddique, M. F., Umar, M., & Kim, J.-M. (2025). Acoustic emission-based pipeline leak detection and size identification using a customized one-dimensional DenseNet. Sensors, 25(4), 1112.

[7] Lu, W., Ye, C., Zhao, C., Wang, J., & Zhao, H. (2023). Leakage identification for mineral air supply pipeline system based on joint noise reduction and ELM. Measurement, 219, 113304.

[8] Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2021). Understanding deep learning (still) requires rethinking generalization. Communications of the ACM, 64(3), 107-115.

[9] Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: Theory and applications. Neurocomputing, 70(1-3), 489-501.

[10] Cambria, E., Huang, G.-B., Kasun, L. L. C., Zhou, H., Vong, C. M., Lin, J., Yin, J., Cai, Z., Liu, Q., Li, K., et al. (2013). Extreme learning machines. IEEE Intelligent Systems, 28(6), 30-59.

[11] Zhu, Q.-Y., Qin, A. K., Suganthan, P. N., & Huang, G.-B. (2005). Evolutionary extreme learning machine. Pattern Recognition, 38(10), 1759-1763.

[12] Miche, Y., Van Heeswijk, M., Bas, P., Simula, O., & Lendasse, A. (2011). TROP-ELM: A double-regularized ELM using LARS and Tikhonov regularization. Neurocomputing, 74(16), 2413-2421.

[13] Miche, Y., Sorjamaa, A., Bas, P., Simula, O., Jutten, C., & Lendasse, A. (2009). OP-ELM: Optimally pruned extreme learning machine. IEEE Transactions on Neural Networks, 21(1), 158-162.

[14] Zhou, S., Liu, X., Liu, Q., Wang, S., Zhu, C., & Yin, J. (2016). Random Fourier extreme learning machine with L2,1-norm regularization. Neurocomputing, 174, 143-153.

[15] Li, L., Han, Q., Li, J., & Cui, Z. (2023). Two-step multi-view data classification based on dynamic graph-ELM. Pattern Recognition Letters, 176, 236-243.

[16] He, X., & Niyogi, P. (2003). Locality preserving projections. Advances in Neural Information Processing Systems, 16.

[17] Wang, W., & Chen, H. (2022). Predicting miRNA-disease associations based on graph attention networks and dual Laplacian regularized least squares. Briefings in Bioinformatics, 23(5), bbac292.

[18] Ding, Y., Tang, J., & Guo, F. (2020). Identification of drug-target interactions via dual Laplacian regularized least squares with multiple kernel fusion. Knowledge-Based Systems, 204, 106254.

[19] Lu, G.-F., Yu, Q.-R., Wang, Y., & Tang, G. (2020). Hyper-Laplacian regularized multi-view subspace clustering with low-rank tensor constraint. Neural Networks, 125, 214-223.

[20] Adeli, E., Thung, K.-H., An, L., Wu, G., Shi, F., Wang, T., & Shen, D. (2018). Semi-supervised discriminative classification robust to sample-outliers and feature-noises. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), 515-522.

[21] Zhang, Y., Zhang, W., Li, Y., Wen, L., & Sun, X. (2024). AFOS-ELM-MVE: A new online sequential extreme learning machine of dam safety monitoring model for structure deformation estimation. Advanced Engineering Informatics, 60, 102345.

[22] Nie, F., Zhang, R., & Li, X. (2017). A generalized power iteration method for solving quadratic problem on the Stiefel manifold. Science China Information Sciences, 60, 1-10.

[23] Chen, Z., Liu, Y., Zhang, Y., Li, Q., & Alzheimer's Disease Neuroimaging Initiative. (2023). Orthogonal latent space learning with feature weighting and graph learning for multimodal Alzheimer's disease diagnosis. Medical Image Analysis, 84, 102698.

[24] Yang, M., Deng, C., & Nie, F. (2019). Adaptive-weighting discriminative regression for multi-view classification. Pattern Recognition, 88, 236-245.

[25] Shao, W., Xiang, S., Zhang, Z., Huang, K., & Zhang, J. (2021). Hyper-graph based sparse canonical correlation analysis for the diagnosis of Alzheimer's disease from multi-dimensional genomic data. Methods, 189, 86-94.

[26] Wang, Q., Guo, Y., Wang, J., Luo, X., & Kong, X. (2018). Multi-view analysis dictionary learning for image classification. IEEE Access, 6, 20174-20183.

[27] Lin, X., Li, G., Wang, Y., Zeng, K., Yang, W., & Wang, F. (2024). Advances in intelligent identification of fiber-optic vibration signals in oil and gas pipelines. Journal of Pipeline Science and Engineering, 4(4), 100184.

[28] Gong, Y., Bao, C., He, Z., Jian, Y., Wang, X., Huang, H., & Song, X. (2025). A review on gas pipeline leak detection: Acoustic-based, OGI-based, and multimodal fusion methods. Information, 16(9), 731.