Research On Fault Identification Methods Based On MRMR And MKELM

Jingzong Yang

School of Big Data, Baoshan University, BaoShan, Yunnan, 678000, P.R.China

Corresponding author. E-mail: yjingzong@foxmail.com

Received: Sep. 29, 2023; Accepted: Mar. 04, 2024

As an emerging detection technology, acoustic active detection has attracted considerable attention in recent years due to its advantages such as low cost, non-destructiveness, and ease of signal collection. To address the challenges of fault identification in acoustic detection of pipeline blockages, this paper proposes a fault recognition method that combines Minimum Redundancy Maximum Relevance (MRMR) and Multi-Kernel Extreme Learning Machine (MKELM). Firstly, statistical analysis and multi-resolution wavelet transform are employed to extract the acoustic pulse response signals of the blockage, thus constructing mixed-domain features. Subsequently, the MRMR algorithm is utilized for secondary feature extraction to eliminate redundant features. Finally, the optimized features are input into the MKELM classification model for identification. Experimental results demonstrate that mixed-domain features are more effective in characterizing blockage states compared to single-domain features, while the MRMR algorithm significantly reduces feature redundancy. Additionally, when comparing the recognition performance of different classification models, MKELM achieves higher accuracy than single-kernel models and traditional methods.

Keywords: Pipeline; MRMR; MKELM; Pattern recognition

http://dx.doi.org/10.6180/jase.202503_28(3).0013

1. Introduction

Pipeline transportation, as a novel approach for long-distance conveyance, offers not only a substantial capacity but also ensures energy efficiency, environmental sustainability, minimal loss, and exceptional security and stability. Notably, in the past decade, long-distance water pipelines, serving as a crucial component of urban infrastructure, have gained increasing significance in society. Concurrently, the issue of pipeline transportation safety has become increasingly pressing. The phenomenon of sediment deposition in the pipeline is very common, which is caused by the poor quality of pipeline construction, or the failure of fluid flow in meeting the design requirements. In this case, if the protection measures were not taken in time, the blocked area will gradually expand and eventually destroy the normal operation of pipeline. Therefore, it is very important to carry out the pipeline blockage detection in time to ensure the safe operation of the pipeline.

There are numerous pipeline blockage detection methods, typically including the percussion sound method, the gamma ray detection method, the pipeline strain method, the artificial neural network method, and the acoustic impulse response analysis method. According to the results of previous research, the detection technologies can be divided into two categories: direct detection and indirect detection. Generally, the direct detection method may have some shortcomings, such as larger errors, dependence on human experience and slowness in response. The indirect detection method not only can be used for qualitative analysis, but also has a great advantage in precise quantitative analysis. Through in-depth analysis of the blocked signals collected by indirect detection method, researchers can obtain the specific information about pipeline blockage. So, this method is a research hotspot in recent years, especially in the analysis of acoustic impulse response based on acoustic detection, which has high sensitivity without causing any damage to pipeline equipment and fast detection speed [1]. It has also been widely used in petroleum and natural gas industry. However, the collected acoustic signal of pipeline is nonlinear and non-stationary, and the characteristics of time domain or frequency domain analysis can only reflect the equipment failure in limited ways. In this regard, it will affect the accuracy of pattern recognition. In the time domain, it reveals the trend and characteristics of the signal over time. In the frequency domain, it represents the distribution signal energy or power across different frequencies. In the time-frequency domain, it allows for the study of varied information across different frequency ranges. Therefore, to fully capture the operational state of the pipeline, it is essential to combine the time domain, frequency domain, and time-frequency domain into a mixed domain feature set. This approach allows for a more comprehensive representation of fault states, capturing a wider range of information. However, the mixed domain feature dataset tends to be quite large, which subsequently increases the demand for time and computational resources. Additionally, the presence of numerous irrelevant and redundant features can complicate the learning algorithm and negatively impact pattern recognition performance.

To enhance computational efficiency and eliminate redundant features, techniques such as Principal Component Analysis (PCA) [2], Independent Component Analysis (ICA) [3], and Partial Least Squares (PLS) [4] have been introduced. By utilizing these algorithms, the dimensionality of features can be rapidly reduced. However, these methods may not yield satisfactory results for the dimensionality reduction of nonlinear and non-stationary data. Based on the correlation and redundancy of features, a new nonlinear analysis method, which is called minimal redundancy maximal relevance (MRMR) is proposed [5]. This algorithm uses mutual information to determine the correlation and redundancy of features. Also, by looking for the feature subset to reduce the dimensionality of the feature, the efficiency and accuracy of classification can be improved [6]. Currently, this algorithm has gained widespread application in the field of feature dimension reduction [7–9].

The primary objective of feature extraction is to facilitate classification. Among the commonly used pattern recognition methods, BP neural network often requires a large amount of data and may easily converge to local optima, posing challenges in terms of practical applications. Extreme Learning Machine (ELM) [10, 11], a relatively recent development, offers advantages in terms of simple parameter selection and high efficiency. However, the input weights and hidden layer thresholds are randomly distributed, leading to variations in classification accuracy. To address these issues, Huang et al. [12] introduced the concept of nuclear mapping in lieu of the traditional random mapping method, resulting in the development of Kernel Extreme Learning Machine (KELM). This method eliminates the need for assigning connection weights between the input and hidden layers, thereby enhancing stability compared to traditional ELM techniques. Meanwhile, the ability of this method in solving the problem of classification and regression forecasting is stronger. However, the choice of kernel function for KELM has a great influence on its performance, and each kernel function has some limitations. Especially when the training sample size is large or the distribution of high-dimensional data is uneven [13–15], the single kernel method cannot achieve good results. So, it is an inevitable trend to construct a classification model which is composed of multiple kernel functions [16, 17], and it can play the advantages of each single kernel function. As a better flexibility model, multiple kernel model can enhance the interpretability of decision function. Compared to single kernel models, Multi-Kernel Extreme Learning Machines (MKELMs) require the determination of more parameters, which can significantly impact their performance. Therefore, the optimization of the combination coefficient in kernel function is a key problem in the construction of multiple kernel feature space. Usually, it is a feasible way to optimize the kernel function parameters by introducing the intelligent optimization algorithm. In recent years, the Shuffled Frog Leaping Algorithm (SFLA) has been introduced as a novel metaheuristic search algorithm based on a collaborative group of swarm intelligence algorithms. This algorithm effectively combines the particle swarm algorithm and memetic algorithm, resulting in high computational efficiency and excellent optimization performance [18]. Meanwhile, by comparing the performance of five kinds of intelligent algorithms with several standard Benchmark functions, the experimental results of Elbeltagi et al [19] also show that the convergence speed and optimization performance of SFLA is good. At present, SFLA has been applied in the optimization problem of path planning, shop scheduling, water quality parameter prediction, etc. [20–22].

Combining fault features with classification methods can be an effective approach to addressing the issue of pipeline blockage. Based on this, this paper proposes a pattern recognition method utilizing MKELM with minimal redundancy maximal relevance (MRMR) feature selection. The primary contributions of this paper are reflected in the

Table 1. Feature Information.

Time domain	Frequency domain
Peak value, Mean value, Root mean square value, Variance, Standard deviation, Peak factor, Kurtosis coefficient, Impulse factor, Margin coefficient, Waveform factor	Gravity frequency, Mean square frequency, Frequency variance, Root mean square frequency, Frequency standard deviation

following two aspects:

1. This paper innovatively employs a secondary feature extraction method based on MRMR, successfully constructing a comprehensive set of sensitive features that integrate the time domain, frequency domain, and time-frequency domain. This approach not only significantly reduces feature redundancy but also provides a more precise and effective data foundation for subsequent pattern recognition of blockage faults.

2. To further enhance the generalization performance and classification accuracy of the model, this paper carefully selects multiple kernel functions and optimizes the MKELM model by integrating the SFLA algorithm. Finally, we successfully apply this optimized model to the identification of pipeline blockage signals, providing new solutions for practical engineering applications.

2. Construction of mixed domain feature sets and feature extraction

2.1. Construction of mixed domain feature sets

In traditional fault diagnosis, to recognize the faults, it is necessary to extracting the corresponding features of the collected signal. However, feature parameters extracted from a single domain only provide limited information, and such features usually reflect the characteristics of faults in limited ways. Therefore, the method, which integrates multiple feature information from the time domain, frequency domain, and time-frequency domain, can effectively combine extensive object information, thereby enabing a more precise and efficient identification of faults.

2.1.1. Construction of time domain and frequency domain features

Time domain features can be extracted through statistical analysis of the signal's probability density function. Ten feature parameters were selected in this study to represent time domain features, which the aim of capturing the amplitude, energy, and distribution of the time series. Furthermore, given that the frequency domain offers more comprehensive information than the time domain in pattern recognition, five commonly used features were selected to comprise the feature set for the frequency domain. The extracted features are summarized in Table 1. Their calculation methods are described in references [23].

2.1.2. Time-frequency domain feature construction based on wavelet decomposition

As the blocked area in the pipeline increases, the acoustic signal undergoes corresponding changes, with transient or mutational signals often containing specific information about the blockage. While single time or frequency domain feature analysis can provide global information, they are limited in their ability to analyze the local properties of the signal. However, as a multi-scale and multi-resolution signal decomposition method, wavelet transform offers advantages over time or frequency domain analysis in processing non-stationary signals. It can fully preserve both high-frequency and low-frequency information from the original signal. Therefore, by incorporating the characteristics specific to pipeline blocking signals, the wavelet transform Mallat algorithm [24] is employed to decompose the signal, followed by the extraction of time-frequency features from the wavelet coefficients. The experiment conducted a three-layer wavelet decomposition on the original acoustic signal of pipeline blockage, which can generate four new time series. Due to the fact that each time series can generate 10 time-domain feature quantities and 5 frequency-domain feature quantities, there are a total of 60 time-frequency domain features.

2.2. Construction of mixed domain feature sets based on MRMR

MRMR utilizes mutual information to assess the correlation and redundancy among features. It employs the information gap and information entropy to formulate the feature subset. Considering the random variables x and y, which possess a joint probability mass function $P(x,y)$ and marginal probability mass functions $P(x)$ and $P(y)$, the mutual information is defined as follows:

$$I(x,y) = \iint P(x,y) \log \frac{P(x,y)}{P(x)P(y)} dxdy \tag{1}$$

The definition of MRMR measure index is as follows:

$$\max D(\mathrm{S},\mathrm{c}), \quad \mathrm{D} = \frac{1}{|S|_{x_i \in S}} \sum^{n} I\left(x_i, c\right) \tag{2}$$

$$\min R(\mathrm{S}), \quad R=\frac{1}{|S|^2} \sum_{x_i, x_j \in S}^{n} I\left(x_i, x_j\right) \tag{3}$$

Where $|S|$ is the dimension of feature space. c is the sample class. $I\left(x_i, c\right)$ is the mutual information between x_i and class c. I is the mutual information between x_i and x_{j}.

To achieve a better classification effect with reduced feature dimension, the essence of this method lies in maximizing the correlation between features and classes while minimizing redundancy among features. Hence, by considering mutual information, the method effectively identifies the relatively optimal feature subset through comparative analysis. Through summing up the above indicators, the criterion of MRMR can be expressed as follows:

$$\max \phi_1(\mathrm{D}, \mathrm{R}), \quad \phi_1=D-R \tag{4}$$

$$\min \phi_2(\mathrm{D}, \mathrm{R}), \quad \phi_2=D / R \tag{5}$$

Generally, different fault types and degrees exhibit varying levels of sensitivity in their characteristics. Additionally, there are significant differences in the state information conveyed by distinct features. Therefore, in our experiment, we initially extract ten-time domain features, five frequency domain features, and sixty time-frequency domain features to construct a comprehensive mixed domain feature set. Subsequently, we employ MRMR for secondary feature extraction, aimed at eliminating redundant data and irrelevant features. Finally, we utilize the selected feature vector to identify the pipeline blocking state. The detailed algorithm flow is outlined as follows:

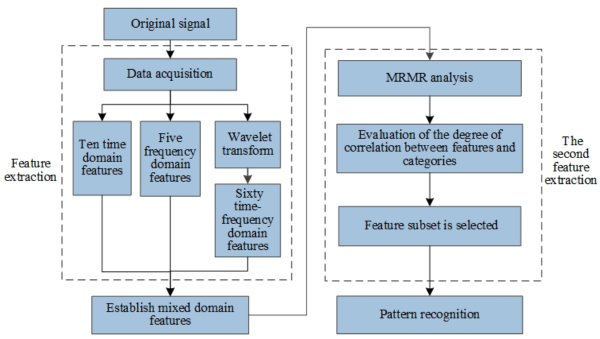


Fig. 1. Flow chart of feature extraction method based on MRMR.

Step 1: Normalize the sample data;

Step 2: Extract the features, and then build the mixed domain features.

Step 3: Calculate the correlation and redundancy between single class feature and target class with MRMR method;

Step 4: Sort the mixed domain features in descending order based on the measurement index obtained from the MRMR method;

Step 5: Select the top q optimal feature as the feature subset based on the sorting results from Step 4;

Step 6: Build a classification model based on extracted feature vectors.

3. Pipeline blockage identification model of mkelm optimized by sfla

3.1. Kernel extreme learning machine

ELM is a network learning algorithm proposed by Huang et al [11]. In the ELM model, the output matrix H of the hidden layer is randomly assigned. This random assignment introduces uncertainty, resulting in different matrix H values in each modeling analysis. This directly affects the stability and generalization capabilities of ELM. To address these issues, Huang et al. [25] proposed the Kernel Extreme Learning Machine (KELM). He applied the Mercer's condition to define the kernel matrix:

$$\left\{\begin{array}{l}\Omega=H H^T \\ \Omega_{i, j}=h\left(x_i\right) \cdot h\left(x_j\right)=K\left(x_i, x_j\right)\end{array}\right. \tag{6}$$

Where Ω is a symmetric matrix of NxN.$K\left(\mathrm{x}_{\mathrm{i}}, \mathrm{x}_{\mathrm{j}}\right)$ is the kernel function. The output of KELM can be obtained as follows:

$$f\left(x_{\mathrm{p}}\right)=\left[K\left(x_{\mathrm{p}}, x_1\right), \ldots, K\left(x_{\mathrm{p}}, x_N\right)\right]\left(\frac{1}{C}+\Omega\right)^{-1} T \tag{7}$$

3.2. Multiple kernel extreme learning machine

The choice of kernel function has a significant impact on the classification. Often relying on experience can significantly limit its generalization performance and classification accuracy. Commonly used kernel functions include the polynomial kernel function (poly) and the radial basis function (RBF). These functions belong to the global kernel functions and local kernel functions, respectively. The RBF kernel can be described as follows:

$$K_{\mathrm{RBF}}\left(x_{\mathrm{p}}, x_i\right)=\exp \left(-\left\|x_{\mathrm{p}}-x_i\right\| / \sigma^2\right) \tag{8}$$

Fig. 2 is a test curve in which the value of σ^2 in RBF kernel function is evaluated to 1, 3, 5 and 7 respectively, and $x_i=0.2$ is used as the test point. It can be seen from Fig. 2 that the value of the kernel function is changed only in a small neighborhood near the test point 0.2, so the RBF kernel is a local kernel function. Meanwhile, The Poly kernel can be described as follows

$$K_{\text{Poly}}\left(x_{\mathrm{p}}, x_i\right)=\left(x_{\mathrm{p}} \cdot x_i+1\right)^d \tag{9}$$

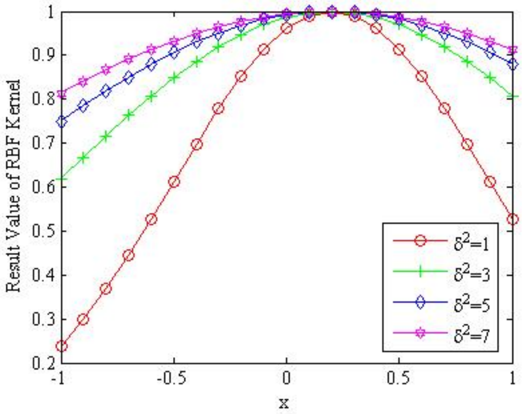


Fig. 2. Test curve of RBF kernel.

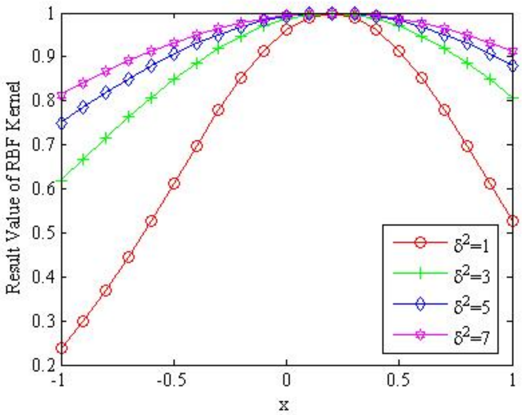


Fig. 3. Test curve of Polynomial kernel.

Fig. 3 is a test curve in which the value of d in Poly kernel function is evaluated to 1,3 , 5 and 7 respectively, and $x_i = 0.2$ is used as the test point. It can be seen from Fig. 3 that the value of the kernel is obviously changed when it is away from the test point 0.2 . Therefore, the polynomial kernel is a global kernel function.

In order to make full use of the advantages of various kernel functions, and reduce the influence of different types of kernel function on the model performance, this paper established the following multiple kernel extreme learning machine (MKELM) model:

$$K_{\text{mulit}}\left(x_{\mathrm{p}}, x_i\right) = \mu K_{\text{RBF}}\left(x_{\mathrm{p}}, x_i\right) + (1-\mu) K_{Pohy}\left(x_{\mathrm{p}}, x_i\right) \tag{10}$$

Where $\mu(\mu \in (0,1))$ is the weight coefficient which used to adjust the two kinds of kernel function. As a result, the output of MKELM is as follows:

$$f\left(x_{\mathrm{p}}\right) = \left[K_{\text{multi}}\left(x_{\mathrm{p}}, x_1\right), \ldots, K_{\text{multi}}\left(x_{\mathrm{p}}, x_{\mathrm{N}}\right)\right] \left(\frac{1}{C} + K_{\text{multi}}\left(x_i, x_j\right)\right)^{-1} T \tag{11}$$

Therefore, the parameters of the training of MKELM model which needs to optimize are $\{C, \sigma^2, d, \mu\}$.

3.3. Parameter optimization of MKELM based on SFLA

After selecting the multiple kernel function, the parameters of MKELM which need to be optimized include the penalty factor C, RBF kernel function σ^2, poly kernel function d and the weighted factor μ. Usually, the algorithm parameters are determined by the experience of the task and a lot of experiments. In this paper, the parameters of MKELM are optimized by the application of intelligent optimization algorithm - shuffled frog leaping algorithm (SFLA).

SFLA [26, 27] simulates the frog's exchange behavior of information within each sub group to find food. The basic idea is to generate an initial population randomly in a D-dimensional space, where the population consists of F frogs. The population is then divided into S sub groups, with each sub population containing n frogs. The division is based on fitness, and the division rules are as follows: the first frog is assigned to the first sub-group, the second frog to the second sub-group, and so on until the Sth frog, which is assigned to the Sth sub-group. The (S+1) th frog is then assigned to the first sub-group, and this process is repeated until all frogs have been assigned to a sub group.

For each sub group of individuals, the best fitness and the worst fitness of the frog is X_b and X_w. Additionally, the best fitness of the population is expressed as X_g. In the sub population of each evolutionary process, do a local search operation of X_w loop, and update it. The detailed strategy is:

$$\mathrm{D_i} = \mathrm{rand}() \cdot (X_{bi} - X_{wi}) \tag{12}$$

$$X_w{}' = X_w + D_i, (-D_{\max} \leq D_i \leq D_{\max}) \tag{13}$$

Where rand (0 is a random number distribution within $[0,1]. D_{\max}$ is the largest step. If the updated frog's fitness $X_w{}'$ is better than the original frog X_w, then the X_w is replaced by $X_w{}'$. If there is no improvement, the X_b is replaced by X_g, and perform local search according to the formula (25) and (26). If there shows no improvement, then generate a new position to replace the original X_w. Repeat the above operation and until the end of the local search. Meanwhile, the whole population is mixed, and then they are divided into groups and sorted to continue the local search again. The above rules are repeated before the completion of the conditions to complete the search. The main process of parameter optimization of MKELM based on SFLA is shown as follows:

Step 1: Normalize the extracted feature vectors, and then divide them into training samples and test samples.

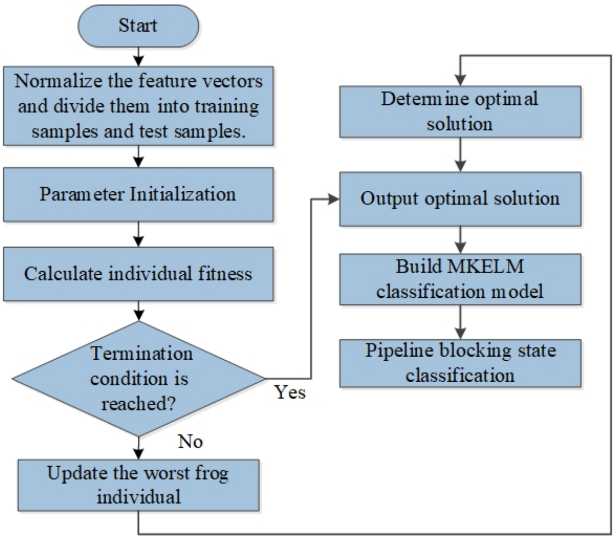


Fig. 4. Flow chart of SFLA-MKELM model.

Step 2: Initialize the parameters. Set the total number of individuals in the population, the number of sub-groups, the number of frogs in each sub-group, the total number of mixed iterations, the total number of iterations in the sub-group, and the range of the four parameters to be searched for MKELM, etc.

Step 3: Calculate the individual fitness. Then divide all individuals into groups. Because the goal of the optimizing MKELM parameters based on SFLA is achieve the maximum classification accuracy, the classification accuracy of sample in MKELM is set as the fitness function. The greater the degree of fitness, the better the solution is optimal.

Step 4: Determine the worst solution of each sub population, the optimal solution of each sub population, and the global optimal solution of the population. And then update the worst frog again and again based on the frog's update strategy for each sub group. At the same time, the worst individuals which have not improved the fitness in sub group are randomly generated by individuals.

Step 5: Mix the updated sub groups when the total number of iteration in sub groups is reached, Then sort the each sub group of individuals according to the size of the fitness value.

Step 6: Check the end condition. When the maximum number of iterations is satisfied, output the optimal solution. Otherwise, go to the Step 2.

Step 7: Put the optimal solution into MKELM model.

4. Experimental analysis

This experiment utilized a 14.2-meter pipeline with an internal diameter of 150 millimeters. To emulate various pipeline blockage scenarios, four distinct test series were conducted for comparative purposes. The initial test series represented the healthy condition of the pipeline, while the subsequent three series simulated varying degrees of blockages placed within the pipeline. These experiments captured blockage signals through acoustic detection methods. The schematic representation of pipeline blockage detection is depicted in Figure 5. The experimental apparatus consisted of a microphone (SPM0208HE5), a loudspeaker (Visaton Type K50WP), a power amplifier, and a computer. The computer, equipped with WinMLS software, functioned as the central controller for the sound card, generating a 10-second sinusoidal sweep. When the sound wave encountered blocking materials, it would reflect and transmit back to the head end. The microphone captured the reflected pulse signals, which were then recorded. The sampling frequency of the signal was set to 44100 Hz. Within these experiments, blockage scenarios were categorized based on their height: 20 millimeters, 40 millimeters, and 55 millimeters for single-point blockages, labeled as Level I, Level II, and Level III blockages, respectively.

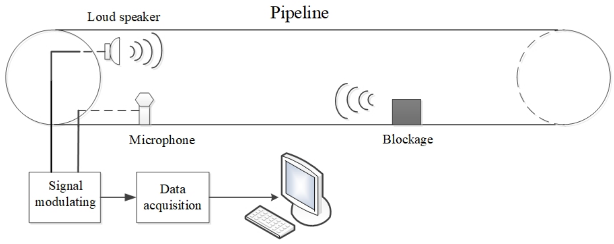


Fig. 5. Schematic diagram of pipeline blockage detection.

The acoustic impulse response of a pipeline is a frequently utilized physical quantity in various studies [29]. Its significance lies in the fact that it encapsulates a plethora of information regarding defects in the pipeline, such as blockages and leakages, making it an invaluable tool in pipeline integrity assessment [30-31]. To facilitate a more nuanced analysis of the detection signal, the signal collected during the experiment undergoes preprocessing. This preprocessing step enables the calculation of the acoustic impulse response of the pipeline. The impulse response is derived through the deconvolution of the original signal, which effectively separates the excitation signal from the system's response, providing a clearer representation of the pipeline's characteristics.

$$h(t) = \mathrm{Re}\left\{ F^{-1} \left[F\left\{ y_i(t) / (f\{x(t)\} + \gamma) \right] \right\} \right. \tag{14}$$

Where $x(t)$ is the excitation signal, $y(t)$ is the signal recorded on microphone.

The time domain diagrams of the four distinct impulse response signals are presented in Fig. 4. These diagrams exhibit the impulse response signals corresponding to various levels of blockage in the pipeline, with labels (a), (b), (c), and (d) denoting the signals for Level I, Level II, Level III, and Level IV blockages, respectively. A comprehensive analysis of Figure 6 reveals that the signal patterns for these four states exhibit a striking similarity, making it challenging to differentiate between the presence or absence of a blockage, as well as to determine the extent of the blockage, solely based on the time domain diagrams. Furthermore, to demonstrate the efficacy and reliability of the method, this paper conducted three comparative experiments.

Experiment one: According to the method mentioned in this paper, we used the statistical analysis and multi-resolution wavelet transform method to extract the mixed domain features of pipeline under different blocking conditions. There are 40 groups of extracted data in health state and other degrees of blockage respectively. Of the 40 groups, the first 28 groups were used as training samples, while the remaining 12 groups were utilized as test samples. To compare the effectiveness of different feature types on classification accuracy, we fed time domain features (10), frequency domain features (5), time-frequency domain features (60), and a total of 75 initial mixed domain features into the KELM model with an RBF kernel for recognition The classification results are presented in Table 2. The labels A, B, C, and D represent the classification results based on time domain feature, frequency domain feature, time-frequency domain feature, and initial mixed domain feature, respectively.

Table 2. Comparative analysis of the classification results.

State	Accuracy/%			
	A	B	C	D
Health	50	33.33	83.33	100
Level I	66.67	66.67	91.67	83.33
Level II	83.33	33.33	66.67	75
Level III	83.33	75	75	83.33
Average value	70.83	52.08	79.17	85.41

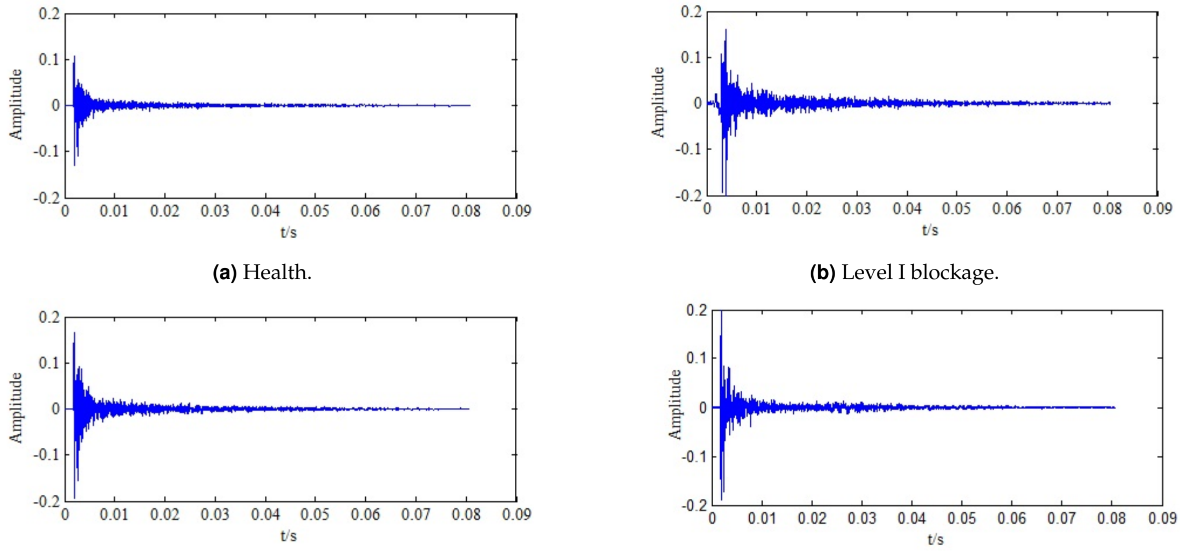
Table 2 reveals that the average classification accuracy achieved using time domain or frequency domain features is relatively low, with values of 70.83% and 52.08%, respectively. However, the average classification accuracy has improved to a certain extent when using time-frequency domain and initial mixed domain feature vectors, achieving values of 79.17% and 85.41%, respectively. It is worth noting that the existence of significant similarity among the initial features does not yield significant advantages in classification using initial mixed domain features. Therefore, further efforts are required to reduce redundant data.

Experiment two: Using the feature dimension reduction method of MRMR, time domain features (10), frequency domain features (5), time-frequency domain features (60), and mixed domain features (75) were analysed. To facilitate their sorting and identification, the features were numbered consecutively. The time domain features, comprising peak value, mean value, root mean square value, variance, standard deviation, peak factor, kurtosis coefficient, impulse factor, margin coefficient, and waveform factor, were numbered from 1 to 10. The frequency domain features, including gravity frequency, mean square frequency, frequency variance, root mean square frequency, and frequency standard deviation, were numbered from 11 to 15. The remaining time-frequency domain features were numbered from 16 to 75.

After optimizing the feature subset using MRMR, feature subsets based on different domains were sorted in descending order based on their quality. Due to space limitations, only the top 10 time-domain features, top 5 frequency features, top 30 time-frequency domain features, and top 30 mixed domain features were selected for presentation. The sorted results are shown in Table 3.

Based on the results presented in Table 3, feature subsets based on different domains were individually input to KELM (RBF kernel) for state recognition. This process involved adding a feature to be identified at each time point. Subsequently, the classification accuracy rate was calculated. The results are depicted in Fig. 7. It can be concluded that when the number of features is gradually increasing, the classification accuracy based on different domains feature subsets are gradually increased. However, it would remain stable or decreased after a peak value. Therefore, based on the analysis results of this test, the top 5, 4, 9 and 15 optimal features extracted from time domain, frequency domain, time-frequency domain and mixed domain were input to KELM respectively for further comparison. Meanwhile, the results of category labels are marked A2, B2, C2 and D2. The accuracy of the classification results is shown in Table 4.

To compare the classification effects, BP neural network and ELM were chosen for testing. The classification accuracy was measured, and the test results were subsequently analyzed. In this test, features were extracted using MRMR from the time domain, frequency domain, time-frequency domain, and mixed domain. These features were then input into the respective classification models. As shown in Fig. 8, the horizontal axis labels T, F, TF, and MIX represent the recognition results based on time domain, frequency domain, time-frequency domain, and mixed domain features,



(a) Health. **(b)** Level I blockage.

(c) Level II blockage. **(d)** Level III blockage

Fig. 6. Pipeline impulse response signal.

Table 3. Comparative analysis of the classification results.

Feature field	Feature ranking
Time domain	7, 1, 10, 8, 9, 6, 2, 4, 5, 3
Frequency domain	11, 12, 15, 13, 14
Time- Frequency domain	22, 16, 17, 37, 18, 52, 19, 20, 51, 24, 21, 25, 26, 27, 28, 53, 29, 23, 67, 30, 31, 33, 34, 36, 35, 39, 40, 41, 68, 42
Mixed domain	22, 1, 10, 7, 11, 37, 12, 52, 13, 14, 8, 21, 15, 51, 16, 17, 18, 19, 20, 24, 23, 53, 25, 67, 26, 27, 28, 6, 29, 30

Table 4. Comparison of classification accuracy.

	Accuracy /%			
State	A_2	B_2	C_2	D_2
Health	58.33	41.67	83.33	100
Level I blockage	83.33	58.33	75	83.33
Level II blockage	66.67	41.67	83.33	91.67
Level III blockage	75	66.67	91.67	83.33
Average value	70.83	52.08	83.33	89.58

respectively.

Upon comparing the classification results of three different models, it is evident that the classification accuracy achieved using mixed domain features is higher than that obtained using single domain features, regardless of the chosen model. Furthermore, the classification results obtained by the KELM model are relatively promising. The KELM utilizes a kernel function to project the input data into a high-dimensional kernel space, effectively mitigating the influence of random fluctuations in the model output associated with ELM. This simplified setting process of hidden layer nodes enhances the model's generalization ability, making it superior to ELM. On the other hand, BP neural networks are prone to overlearning during the training process, resulting in relatively lower classification accuracy compared to other models.

To compare the advantages of MRMR, this study also utilized the feature dimension reduction method of principal component analysis (PCA), which is commonly used to reduce the dimensions of feature subsets based on different domains. In engineering applications, when the cumulative contribution rate exceeds 90%, it indicates that the principal components effectively summarize most of the original data. Therefore, the experiment used this criterion as the basis for reducing the feature dimensions. The result of PCA feature extraction is presented in Table 5. In the dimension reduction analysis of time domain features, it was found that the contribution rate reached 92.81% when using 2 principal components. For frequency domain features, the contribution rate was 99.96% with 1 principal component. For time-frequency domain features, the contribution rate was 93.07% with 5 principal components. Finally, for mixed domain features, the contribution rate was 93.23% with 6 principal components. So, through the

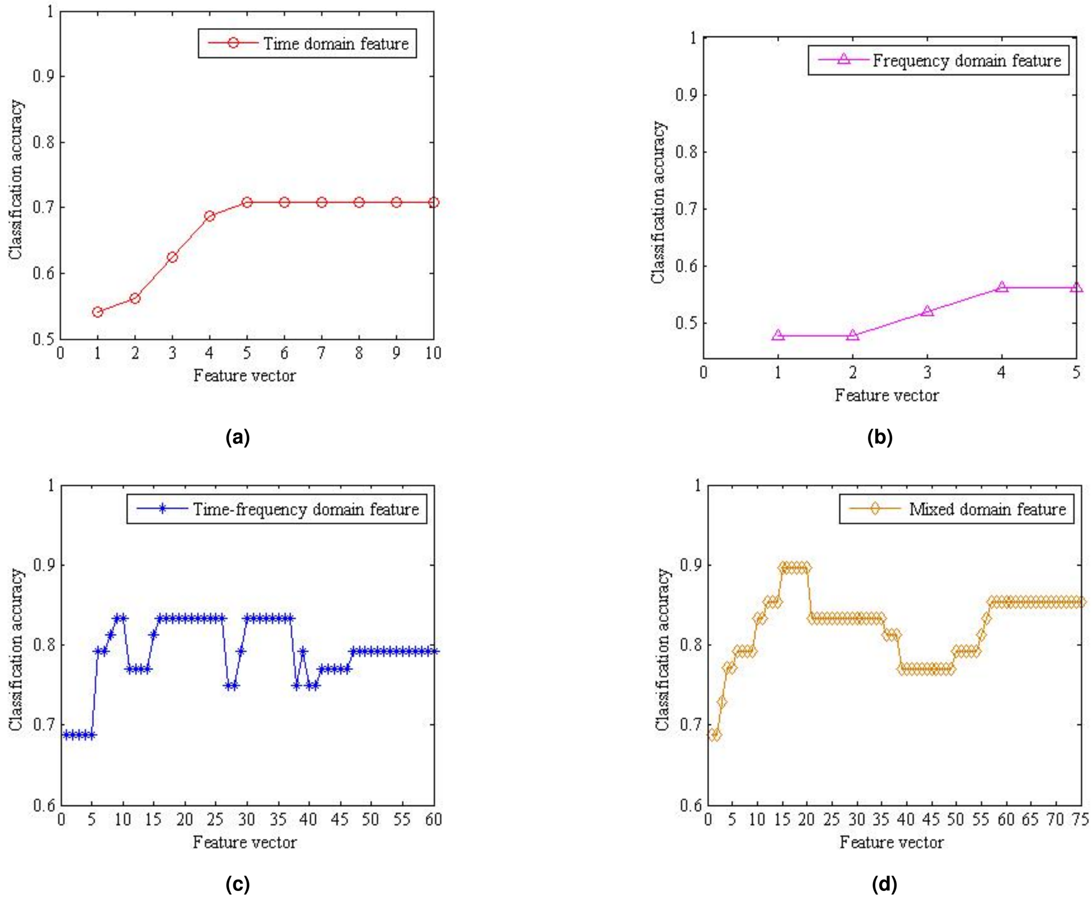


Fig. 7. Classification results of different domains with the change of the number of features.

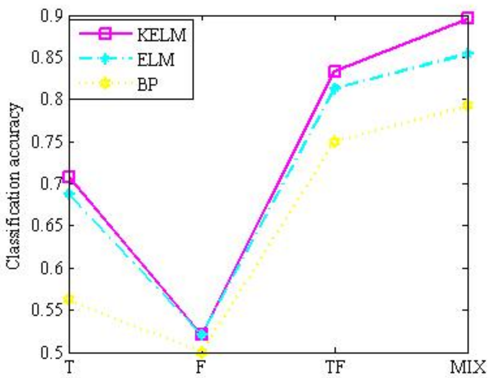


Fig. 8. Recognition results with MRMR.

feature dimension reduction of PCA, 2 time domain features, 1 frequency domain features, 5 time-frequency domain features and 6 mixed domain features are extracted as feature subsets respectively. Then the above four kinds of features were input to KELM for classification, and the result categories are labeled Ak, Bk,Ck and Dk respectively. The classification results are shown in Table 6.

Table 5. Analysis results of PCA feature extraction.

	PCA			
	T	F	TF	MIX
Principal components	2	1	5	6
Contribution rate /%	92.81	99.96	93.07	93.23

Table 6. Comparison of classification accuracy of based on PCA.

	Accuracy /%			
State	A_k	B_k	C_k	D_k
Health	83.33	41.67	66.67	83.33
Level I blockage	66.67	66.67	100	100
Level II blockage	75	33.33	66.67	75
Level III blockage	58.33	58.33	91.67	83.33
Average value	70.83	50	81.25	85.41

Fig. 6 demonstrates that, following PCA-based feature extraction, the classification accuracy achieved using mixed domain features is 85.41%. This accuracy surpasses the classification accuracies obtained using features based on time domain, frequency domain, and time-frequency domain. This indicates that mixed domain features offer a more comprehensive representation of fault characteristics. While PCA effectively reduces the dimensionality of input features, it is evident from Tables 4 and 6 that the average classification accuracy achieved using MRMR feature extraction outperforms that of PCA. Specifically, in the case of mixed domain features, the average classification accuracy with PCA dimension reduction is 85.41%, whereas the average classification accuracy with MRMR dimension reduction is 89.58

To compare the recognition effects of various models using different domain features, we chose the BP neural network and ELM for testing classification accuracy. Additionally, we input different domains feature subsets extracted by PCA into these classification models. As shown in Fig. 10, the labels T, F, TF, and MIX along the horizontal axis represent the recognition results based on time domain, frequency domain, time-frequency domain, and mixed domain features, respectively. Based on Fig. 9, it can be observed that classification results using mixed domain features are superior to those using single domain features within the same classification model. Furthermore, the classification results of KELM outperform other models. By comparing the results of Figs. 8 and 9, it is evident that after feature reduction using MRMR or PCA, the redundancy between features is eliminated, leading to an overall improvement in the classification accuracy of the model. Notably, the classification model based on mixed domain features achieves the highest accuracy when compared to other single domain features.

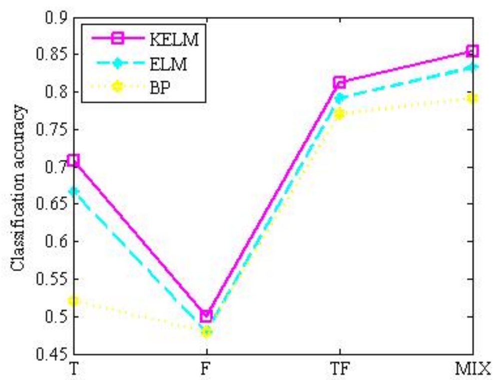


Fig. 9. Recognition results with PCA.

Experiment three: to compare the testing result of different kernel functions, the KELM (Poly kernel), KELM (RBF kernel) and MKELM are optimized by the SFLA algorithm to identify the pipeline blocking signal. The training samples and testing samples are identical in each classification model, and the feature vectors which were input to the classification models are mixed domain features based on MRMR analysis. The initialization parameters are as follows: the total number of frogs is set to 100, the number of sub-groups is 10, the number of frogs within each sub-group is 10, the number of iterations within each sub-group is 10, and the total number of mixed iterations is 50. Meanwhile, the parameter search scope of penalty factor C, RBF kernel function σ^2, poly kernel function d and weighted coefficient μ are: $C \in [0, 1500], \sigma^2 \in [0, 0.5], d \in [0, 10]$ and $\mu \in [0, 1]$ respectively. Then the classification model is built based on the initial parameters, and the parameters in this model are optimized by SFLA. Fig. 10 shows the fitness value transformation curve of a SFLA optimized multi kernel classification model. After the optimization of SFLA, the optimal parameters and classification accuracy of each model are shown in Table 7. Taking into account the randomness of the initialization of the SFLA, the experiment used the classification model based on Poly kernel, RBF kernel and the multiple kernels to optimize for several times. Table 7 lists some of the results.

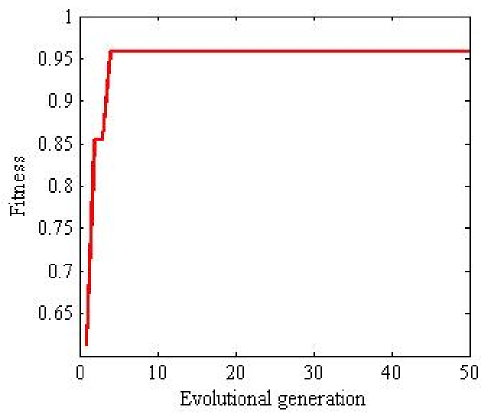


Fig. 10. Fitness value.

As shown in Table 7, the classification accuracy based on the multiple kernel function is higher than that of the classification model with single kernel function. In addition, for the classification model with multiple kernel function, the optimal parameter values are not limited to a certain range, indicating the flexibility and adaptability of the model. This flexibility results in excellent classification accuracy and generalization performance.

In summary, the three experiments demonstrate that the feature extraction using MRMR generally enhances the

Table 7. Classification results based on different kernel function.

Type of kernel function	weighted coefficient	Accuracy (%)
Poly Kernel	/	89.58
Poly Kernel	/	89.58
Poly Kernel	/	89.58
RBF Kernel	/	91.67
RBF Kernel	/	89.58
RBF Kernel	/	91.67
Multi-Kernel	0.960	93.75
Multi-Kernel	0.624	95.83
Multi-Kernel	0.859	95.83

classification accuracy. Furthermore, the mixed domain feature is more effective than the single domain feature in capturing the blocking state of the pipeline. Additionally, the MKELM classification model, optimized using SFLA, exhibits superior performance compared to other single kernel classification models in terms of accuracy.

5. Conclusions

To address the issue of blocking state recognition, this study employed the methods of MRMR and MKELM to identify the blocking state of pipelines. The following conclusions were drawn:

1. The initial mixed domain feature can be obtained through statistical analysis and multi-resolution wavelet decomposition. In this study, ten-time domain features, five frequency domain features, and sixty time-frequency domain features were considered as initial mixed domain features. Compared to single domain features, the mixed domain feature is more effective in identifying pipeline blocking state information.

2. After extracting features with high sensitivity using MRMR, the redundancy of data can be effectively eliminated. First, the correlation and redundancy between single class features and the target class are computed. Then, based on the MRMR index, the mixed domain feature is sorted in descending order, and the first 15 feature vectors are selected as the optimal feature subset. Compared to the recognition results of the first feature extraction, the classification accuracy generally improved after the second feature extraction. Especially in the classification result based on mixed domain feature, the average classification accuracy is the highest.

3. By comparing the recognition effects of different models based on mixed domain features, it was found that the accuracy of MKELM is higher than that of single kernel classification models. Therefore, multiple kernel functions exploit the advantages of each individual kernel function. Furthermore, the classification results indicate that the accuracy of MKELM not only surpasses that of single kernel models but also exceeds that of ELM, KELM, and BP classification models.

Acknowledgements

This work was supported by the Yunnan Fundamental Research Projects (NO. 202301AT070256), training Program for Baoshan Xingbao Talents, 10th batches of Baoshan young and middle-aged leaders training project in academic and technical (No. 202109).

References

[1] S. Datta and S. Sarkar, (2016) *"A review on different pipeline fault detection methods"* **Journal of Loss Prevention in the Process Industries *41***: 97–106. DOI: https://doi.org/10.1016/j.jlp.2016.03.010.

[2] J. Liu, H. Kang, W. Tao, H. Li, D. He, L. Ma, H. Tang, S. Wu, K. Yang, and X. Li, (2023) *"A spatial distribution – Principal component analysis (SD-PCA) model to assess pollution of heavy metals in soil"* **Science of The Total Environment *859***: 160112. DOI: https://doi.org/10.1016/j.scitotenv.2022.160112.

[3] G. Lakshmi Priya Palla and A. Kumar Pani, (2023) *"Independent component analysis application for fault detection in process industries: Literature review and an application case study for fault detection in multiphase flow systems"* **Measurement *209***: 112504. DOI: https://doi.org/10.1016/j.measurement.2023.112504.

[4] Y. Liu, L. Li, S. Zhao, and C. Zhou, (2022) *"A reliability analysis method based on adaptive Kriging and partial least squares"* **Probabilistic Engineering Mechanics *70***: 103342. DOI: https://doi.org/10.1016/j.probengmech.2022.103342.

[5] H. Peng, F. Long, and C. Ding, (2005) *"Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy"* **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE *27***(8): 1226–1238. DOI: 10.1109/TPAMI.2005.159.

[6] C. Ding and H. Peng, (2005) *"Minimum redundancy feature selection from microarray gene expression data"* **Journal of bioinformatics and computational biology 3**(02): 185–205.

[7] X. Gu, J. Guo, L. Xiao, and C. Li, (2022) *"Conditional mutual information-based feature selection algorithm for maximal relevance minimal redundancy"* **APPLIED INTELLIGENCE** ***52***(2): 1436–1447. DOI: 10.1007/s10489-021-02412-4.

[8] Z. Ren, G. Ren, and D. Wu, (2022) *"Deep Learning Based Feature Selection Algorithm for Small Targets Based on mRMR"* **MICROMACHINES** ***13***(10): DOI: 10.3390/mi13101765.

[9] N. M. Ali, A. I. B. Farouk, S. I. Haruna, H. Alanazi, M. Adamu, and Y. E. Ibrahim, (2022) *"Feature selection approach for failure mode detection of reinforced concrete bridge columns"* **CASE STUDIES IN CONSTRUCTION MATERIALS** ***17***: DOI: 10.1016/j.cscm.2022.e01383.

[10] J. Wang, S. Lu, S.-H. Wang, and Y.-D. Zhang, (2022) *"A review on extreme learning machine"* **MULTIMEDIA TOOLS AND APPLICATIONS** ***81***(29): 41611–41660. DOI: 10.1007/s11042-021-11007-7.

[11] J. S. Manoharan, (2021) *"Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm"* **Journal of Soft Computing Paradigm (JSCP)** ***3***(02): 83–95.

[12] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, (2012) *"Extreme Learning Machine for Regression and Multiclass Classification"* **IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS PART B-CYBERNETICS** ***42***(2, SI): 513–529. DOI: 10.1109/TSMCB.2011.2168604.

[13] F. R. Bach, (2008) *"Consistency of the group Lasso and multiple kernel learning"* **JOURNAL OF MACHINE LEARNING RESEARCH** ***9***: 1179–1225.

[14] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet. "More efficiency in multiple kernel learning". In: *Proceedings of the 24th international conference on Machine learning*. 2007, 775–782.

[15] C. Cortes, M. Mohri, and A. Rostamizadeh. "Learning sequence kernels". In: *2008 IEEE Workshop on Machine Learning for Signal Processing*. IEEE. 2008, 2–8.

[16] D. P. Lewis, T. Jebara, and W. S. Noble. "Nonstationary kernel combination". In: *Proceedings of the 23rd international conference on Machine learning*. 2006, 553–560.

[17] C. Ong, A. Smola, and R. Williamson, (2005) *"Learning the kernel with hyperkernels"* **JOURNAL OF MACHINE LEARNING RESEARCH** ***6***: 1043–1071.

[18] A. Rahimi-Vahed and A. H. Mirzaei, (2007) *"A hybrid multi-objective shuffled frog-leaping algorithm for a mixed-model assembly line sequencing problem"* **COMPUTERS & INDUSTRIAL ENGINEERING** ***53***(4): 642–666. DOI: 10.1016/j.cie.2007.06.007.

[19] E. Elbeltagi, T. Hegazy, and D. Grierson, (2005) *"Comparison among five evolutionary-based optimization algorithms"* **ADVANCED ENGINEERING INFORMATICS** ***19***(1): 43–53. DOI: 10.1016/j.aei.2005.01.004.

[20] A. M. Dalavi, P. J. Pawar, and T. P. Singh, (2016) *"Tool path planning of hole-making operations in ejector plate of injection mould using modified shuffled frog leaping algorithm"* **Journal of Computational Design and Engineering** ***3***(3): 266–273.

[21] A. Rahimi-Vahed and A. H. Mirzaei, (2008) *"Solving a bi-criteria permutation flow-shop problem using shuffled frog-leaping algorithm"* **SOFT COMPUTING** ***12***(5): 435–452. DOI: 10.1007/s00500-007-0210-y.

[22] N. Mahmoudi, H. Orouji, and E. Fallah-Mehdipour, (2016) *"Integration of Shuffled Frog Leaping Algorithm and Support Vector Regression for Prediction of Water Quality Parameters"* **WATER RESOURCES MANAGEMENT** ***30***(7): 2195–2211. DOI: 10.1007/s11269-016-1280-3.

[23] X. Wen. *Pattern recognition and condition monitoring*. Science Press, 2007.

[24] S. MALLAT, (1989) *"A THEORY FOR MULTIRESOLUTION SIGNAL DECOMPOSITION - THE WAVELET REPRESENTATION"* **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE** ***11***(7): 674–693. DOI: 10.1109/34.192463.

[25] G.-B. Huang, (2014) *"An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels"* **COGNITIVE COMPUTATION** ***6***(3, SI): 376–390. DOI: 10.1007/s12559-014-9255-2.

[26] M. Eusuff, K. Lansey, and F. Pasha, (2006) *"Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization"* **ENGINEERING OPTIMIZATION** ***38***(2): 129–154. DOI: 10.1080/03052150500384759.

[27] B. B. Maaroof, T. A. Rashid, J. M. Abdulla, B. A. Hassan, A. Alsadoon, M. Mohamadi, M. Khishe, and S. Mirjalili, (2022) *"Current Studies and Applications of Shuffled Frog Leaping Algorithm: A Review"* **ARCHIVES OF COMPUTATIONAL METHODS IN ENGINEERING** ***29***(5): 3459–3474. DOI: 10.1007/s11831-021-09707-2.