Mechatronic Systems and Control, Vol. 53, No. 10, 2025
An Open Access Paper

HIERARCHICAL SVM WITH DATA SUPERPOSITION FOR STATISTICAL PROCESS CONTROL PATTERN RECOGNITION

Juanjuan Qin,*,** Baokang Zhang,** Qi Wu,** Ning Sun,* and Maoqi Cao*

Abstract

Statistical process control is an effective quality control method, and its wide application has important practical significance in modern manufacturing industry. This paper studies a support vector machine classifier model based on the superposition method, which is suitable for pattern recognition of control chart of statistical process control and can effectively improve the intelligence level of quality monitoring in production. We also design a data preprocessing method based on sample information overlay to improve the recognition rate of sample features in response to the problem of unclear abnormal features in observed data. Through the comparison between the statistical characteristics and shape characteristics of the sample data, a two-layer support vector machine classifier model is established. Using the algorithm of Fruit fly optimisation to optimise the parameters of the two-layer support vector machine classifier can improve the recognition accuracy of the model. The effectiveness and superiority of the proposed model are verified by simulation experiments.

Key Words

Support vector machine, statistical process control, control chart patterns recognition, superposition method

1. Introduction

Statistical process control (SPC), as an effective quality control method, has been widely used in modern manufacturing process quality control [1]–[4]. The key to implement quality intelligent monitoring of statistical process control is to accurately control chart patterns. This is of great significance for detecting anomalies in production, eliminating hidden trouble and improving product quality and reliability [5]–[8]. In recent years, extensive researches on the recognition of control graph patterns based on artificial neural network control graph patterns of original data, artificial neural network control graph patterns of feature extraction, deep learning control graph patterns and support vector machine control graph patterns [9]–[14], have been conducted. Addeh *et al.* [9] used the association rule method to select shapes and statistical features, and proposed a radial basis function neural network recognition method. El-Midany *et al.* [10] proposed a neural network recognition model for multivariate observation data using X^2 as the input of neural network multi-module structure. Zan *et al.* [12] proposed a generalised neural network recognition model based on original data. However, the original data dimension of control chart is high and the training data is huge. These methods have long training time and low recognition accuracy. To solve the problem of high dimensionality of original data, the feature extraction method is used to reduce the computational dimension of artificial neural network. Gauri and Chakraborty [13] used a systematic method based on classification and regression tree to select features and established an artificial neural network recognition model. The above studies indicate that artificial neural network is an effective intelligent recognition method of control graph. The computational complexity of artificial neural network can be greatly reduced, and the accuracy and rapidity of classification can be improved by selecting suitable features. However, artificial neural network is based on the principle of empirical risk minimisation, which has some problems such as difficult to convergence, easy to fall into local extreme value, and difficult to determine the most suitable network structure. Based on the support vector machine (SVM), they have achieved unsupervised pattern recognition [15], [16]. Deep learning has high classification accuracy for data sets with a large number of samples and does not require manual feature extraction. However, in the case of a small number of samples, the classification

* Department of Electronic and Information Engineering, Bozhou University, Bozhou, Anhui 236800, China; e-mail: 2016020060@bzuu.edu.cn; 2023020014@bzuu.edu.cn; caomaoqi@mail.ustc.edu.cn
** College of Information Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China; e-mail: 2111903169@zjut.edu.cn; qwu@zjut.edu.cn
Corresponding author: Maoqi Cao

Recommended by Xingxing Jiang
(DOI: 10.2316/J.2025.201-0564)

accuracy is low [17], [18]. It is difficult to collect abnormal pattern samples of control chart in actual production on a large scale, so deep learning is difficult to be used for small sample control chart pattern recognition [19]. Current SPC pattern recognition approaches, such as traditional SVM and neural networks, typically assume fixed standard deviations (σ) and abnormal starting points (t_0), are unrealistic in industrial scenarios. This work addresses this gap by proposing a hierarchical SVM model with data superposition, specifically designed for robust classification of patterns under variable σ and random t_0. The main objectives are as follows.

1) Develop a hierarchical classifier to distinguish directional trends (flat/ascending/descending) before mode-specific classification to improve computational efficiency.
2) Integrate data superposition and FOA optimisation to enhance robustness against dynamic anomalies.

The above control chart recognition models have high classification accuracy. However, the samples they often solve are those with abnormal starting point at the origin or fixed points in the control chart window. The standard deviation of each model sample is fixed. There is random disturbance in actual production, which is similar to Gaussian white noise and has contingency and uncertainty. Therefore, the standard deviation of the sample may not be a definite value. In real production, the real-time data collected is entered from the right side of the control chart window. The starting point of abnormal patterns constantly changes, which may lead to low recognition accuracy of the above model in actual detection. To solve these problems, a pattern recognition model of support vector machine control is proposed based on superposition method. The superposition method is used to preprocess the original observation data of the sample to improve the identification of the data in the direction of change. The computation dimension is reduced and the computation speed and recognition accuracy are improved by applying the feature groups to different decision levels in the two-layer classifier module. A two-layer classifier model is established by optimising the parameters of drosophila algorithm [20] to SVM. The innovation of this study lies in the following.

1) Proposing a hierarchical SVM architecture for SPC dynamic anomaly recognition, which distinguishes pattern categories through hierarchical feature development system.
2) Using fruit fly optimisation algorithm to automatically adjust parameters, improve accuracy and adaptability to industrial scenarios with variable process parameters.

2. Research Method

2.1 The Basic Pattern of the Control Chart

There are many types of control chart, as shown in Fig. 1. The most widely used is the mean control chart. In 1958, Western Electric Company of the United States proposed six basic control chart modes, namely normal (NOR), cyclic (CYC), upward trend (UT), upward shift (US), downward trend (DT), and downward shift (DS) modes. The NOR mode indicates that the production process is under statistical control [21]. The data in the window often obey normal distribution with the mean μ and the standard deviation σ. The observed values of each sampling point fall within the upper and lower control limits (UCL, LCL). The other five control chart modes are abnormal modes. Due to abnormal disturbance, the data after the abnormal starting point t_0 presents cycle changes (CYC mode), continuous upwards or downwards changes (UT and DT mode), and sudden upwards and downwards changes (US and DS mode) [22]. It is found that the trend pattern could be caused by wear and tear on key parts of the machine tool. Changes in operators, raw materials, or equipment may cause step patterns. Periodic patterns may be related to voltage fluctuations in the power supply. In production, accurate classification and recognition of the collected samples of each control chart pattern is the premise to determine whether the production process is controlled or not and the key to ensure the production quality of the product.

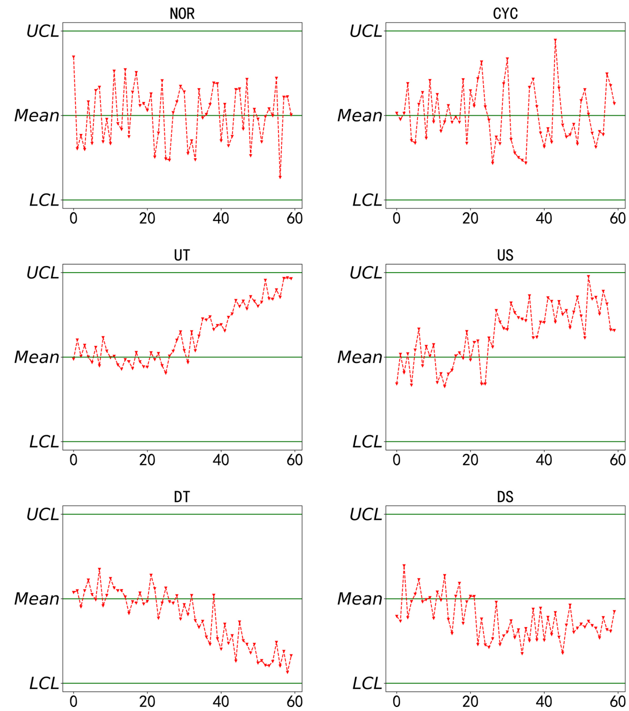

Figure 1. Common control chart patterns.

2.2 Construction of Machine-controlled Mapping Recognition Model Based on Superposition Method

Considering the abnormal starting point and standard deviation change in real production, a support vector machine control mapping pattern recognition model based on superposition method is proposed. It uses Monte Carlo method to generate simulation training and test sample data that are closer to the actual production process. The system structure is shown in Fig. 2. It mainly includes the original data module based on simulation,

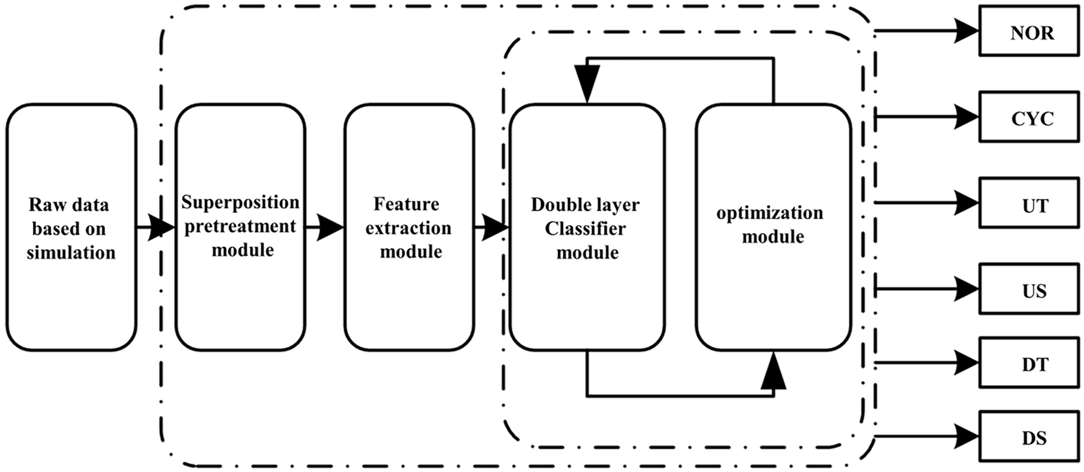

Figure 2. System flowchart.

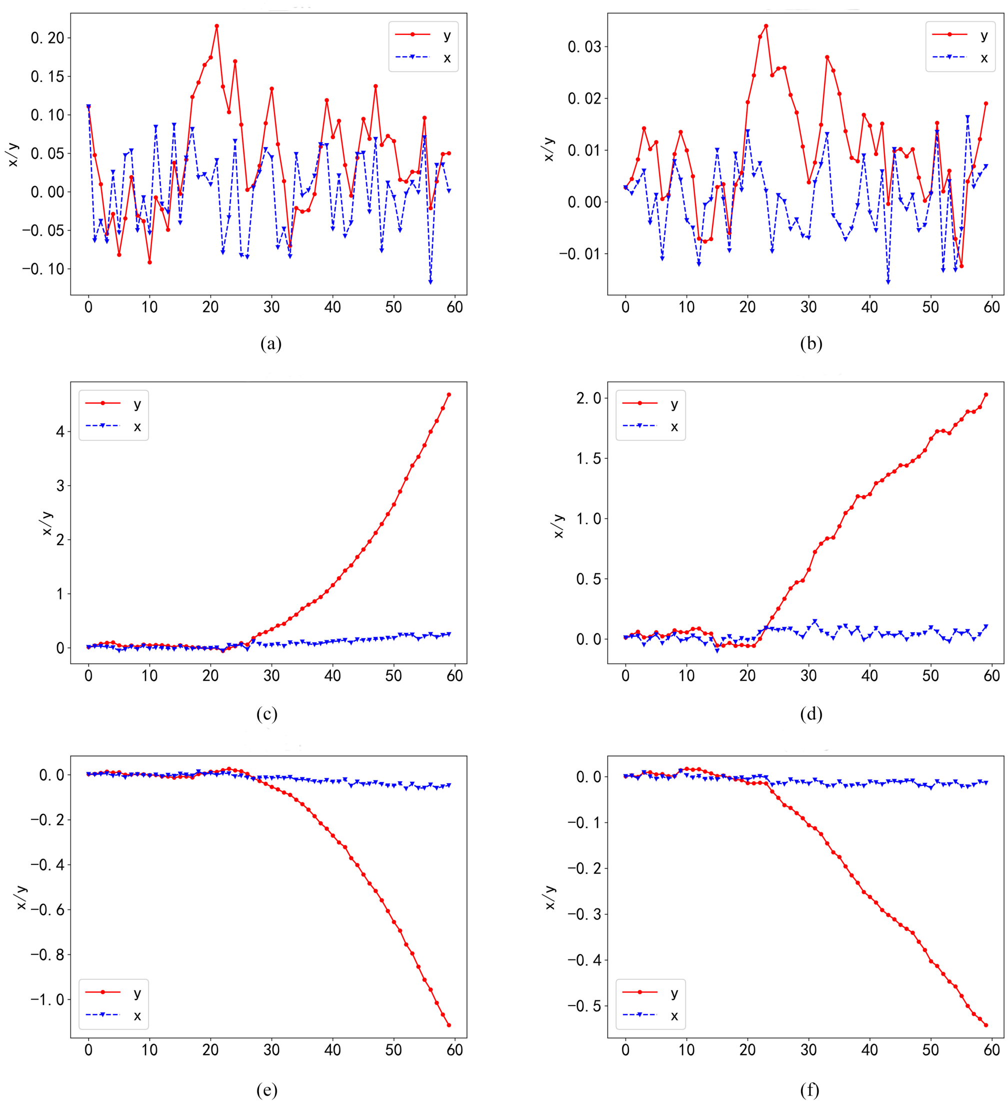

Figure 3. Comparison between the superposed value and original value: (a) NOR; (b) CYC; (c) UT; (d) US; (e) ET; and (f) DS.

superposition pretreatment module, feature extraction module, double layer classifier module and optimisation module. In the superposition pretreatment module, the superposition method is used to preprocess the original data of the control chart. The obtained results can highlight the flatness, rising or falling direction of each model sample to the maximum extent. In the feature extraction module, the calculation dimension is reduced and the speed is improved by extracting statistical and shape features of samples. In the double layer classifier module, a double-layer decision classifier based on vector machine is proposed. The features are divided into two groups for different decision levels in the following two-layer classifier module based on the characteristic of the features. The classification performance of vector machine depends on the penalty coefficient C and the kernel parameter g. As a result, the fruit fly optimisation algorithm is used to optimise the two super parameters C and g of SVM in the optimisation module.

2.3 Superposition Pretreatment

As shown in Fig. 1, abnormal features in the original data of each model sample are not obvious. The amplitude of abnormal data is small, and the identification degree of the sample features is low. The superposition method is used to preprocess the original data, where the original value in each sample is x and the result after superposition method can be obtained as follows:

$$y_n = \sum_{i=1}^{n} x_i \qquad (n-1, 2, \cdots N) \tag{1}$$

where y_n represents the value of each sampling point in the sample after superposition; x_i represents the original value of the current sampling point i in a single sample, and N represents the sampling times of a single sample.

The superposition method is to superpose the observed value corresponding to the current sampling point i and the original value corresponding to the previous sampling point *i-1*. For the samples with basically symmetrical mean values, such as NOR and CYC mode, y_n oscillates along the mean after superposition, and the least squares fitting line tends towards a straight line. For samples with many sampling points whose original values are higher than the mean line, such as UT and US mode, y_n will become larger and larger after the superposition method. Similarly, for DT and DS mode samples, y_n will decrease after superposition. As shown in Fig. 3, the blue and red curves represent the original value and superposed value, respectively. NOR and CYC mode tend to be flat after the superposition method. Abnormal part in the graph becomes steep and the amplitude of the UT, US, DT, and DS modes increase. It can be seen that the horizontal, upward, and downward direction features of the data processed by the superposition method are highlighted. The identification degree of samples can be greatly increased by extracting appropriate statistical and shape features.

2.4 Feature Extraction

Control chart pattern sample has its own statistical characteristics and shape characteristics, which are divided into global shape characteristics and segmented shape characteristics. Suitable features can improve the accuracy and greatly reduce the complexity of the classifier, making

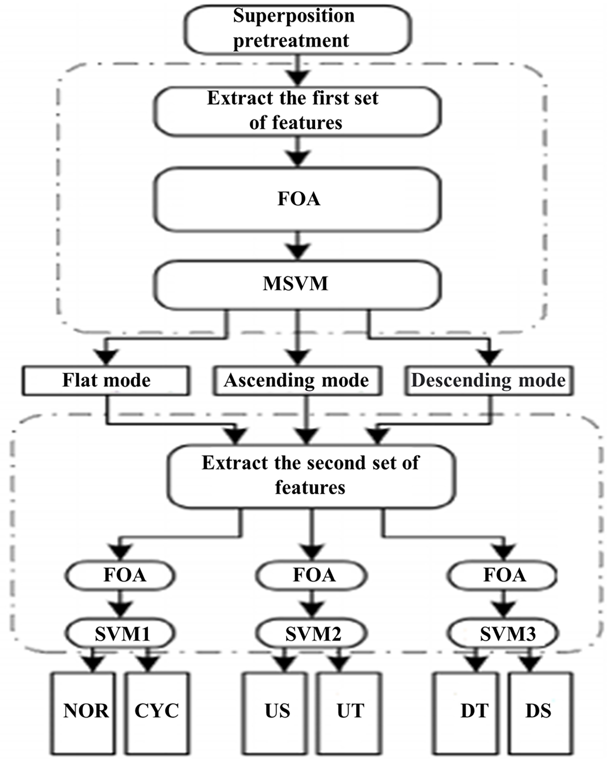

Figure 4. Cognition model.

the classification problem more effective and efficient. Three statistical features: skew, mean variation of late change (Vm), standard deviation variation of late change (VS), six shape features: slope of the line fitted by the least square method (S), proportion of the number of the crossovers to mean line (PMLC), proportion of the number of the crossovers to least square line (PLSC), area between the overall pattern and mean line (ACL), the average slope of the straight lines passing through six pairwise combinations of midpoints of four equal segments (ASL), the ratio of mean sum of squares of errors (MSE) of the least square line fitted to excess data, and average mean sum of squares of errors of the least square line fitted to six subsets of $N/2$ data points (REAE), are extracted. The specific calculation formulas are shown in Table 1.

2.5 Recognition Model of Machine-controlled Mapping-based on Superposition Method

As shown in Fig. 4, the features are divided into two groups according to their distribution characteristics. The first set of features is used to determine the direction of change of the control chart, such as flat mode, ascending mode and descending mode. The second group is used to determine the type of change in the control chart, such as NOR, CYC, US, UT, DT, and DS models. Then, a two-layer classifier model with two main decision levels is designed for two different groups of features.

Decision Level 1 is a multi-classification support vector machine (MSVM) based on superposition method. First, the superposition method is used to process the original observation data. Second, the first set of features, such as Skew, Vm, S, ASL, PMLC and PLSC, are extracted

Table 1
Three Statistical Features and Six Shape Features

Features	Names	Formulas
Statistical features	Skew	$\text{skew} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{N\sigma^3}$
	Mean variation of the late change (Vm)	$V_m = \frac{2}{N}\sum_{i=2/N}^{N} x_i$
	Standard deviation variation of the late change (Vs)	$V_s = \sqrt{\frac{2}{N}\sum_{i=2/N}^{N}(x_i - \overline{x_l})^2} - \sqrt{\frac{2}{N}\sum_{i=1}^{2/N}(x_i - \overline{x_e})^2}$
Global shape features	Slope of the line fitted by the least square method (S)	$S = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(t_i - \bar{t})}{\sum_{i=1}^{N}(t_i - \bar{t})^2}; \bar{t} = \sum_{i=1}^{N}\frac{t_i}{N}$
	• Proportion of the number of the crossovers to mean line (PMLC)	$PMLC = \frac{\sum_{i=1}^{n-1} o_i}{N}; o_i = \begin{cases} 1, (x_i - \overline{x})(x_{i+1} - \overline{x}) \leq 0 \\ 0, (x_i - \overline{x})(x_{i+1} - \overline{x}) > 0 \end{cases};$
	• Proportion of the number of the crossovers to least square line (PLSC)	$PLSC = \frac{\sum_{i=1}^{n-1} {o_i}'}{N}; o'_i = \begin{cases} 1, (x_i - {x_i}')(x_{i+1} - {x_{i+1}}') \leq 0 \\ 0, (x_i - {x_i}')(x_{i+1} - {x_{i+1}}') > 0 \end{cases}$
	Area between the overall pattern and mean line (ACL)	$ACL = \sum_{i=1}^{N-1} \text{Area}$
Segmented shape features	• Average slope of the straight lines passing through six pairwise combinations of midpoints of four equal segments (ASL)	$ASL = \frac{\sum_{\text{jk}} S_{\text{jk}}}{6}, (j = 1,2,3; k = 2,3,4; j < k)$
	• Ratio of mean sum of squares of errors (MSE)of the least square line fitted to over data, and average mean sum of squares of errors of the least square line fitted to six subsets of N/2 data points (REAE)	$REAE = \frac{\text{MSE}}{\sum_{\text{jk}} \text{MSE}_{\text{jk}}/6}; (j = 1,2,3; k = 2,3,4; j < k)$

from the data processed by the superposition method. As shown in Fig. 5, the sample data is divided into flat mode, ascending mode, and descending mode, suggesting that the first group of features has a good discrimination of flat mode, ascending mode, and descending mode. Then, the first set of extracted features is normalised to generate the feature vector.

Decision level 2 consists of three SVM binary classifiers. First, the second set of features is extracted from the sample data of the flat, ascending and descending modes obtained from decision level 1. The second set of characteristics, such as Vs, PMLC, PLSC, ASL, ACL, and REAE, can well distinguish the types of NOR, CYC, US, UT, DT and DS. For example, as shown in Fig. 6, Vs and REAE can effectively distinguish between trend and step patterns. Then, the second set of extracted features is normalised to form the feature vector. Finally, the second group of three mode feature vectors is sent into their respective SVM models for classification and determination. The flat mode samples are divided into normal and periodic modes by SVM1. The ascending mode and descending mode are divided into ascending trend, ascending step, descending trend, and descending step mode by SVM2 and SVM3, respectively.

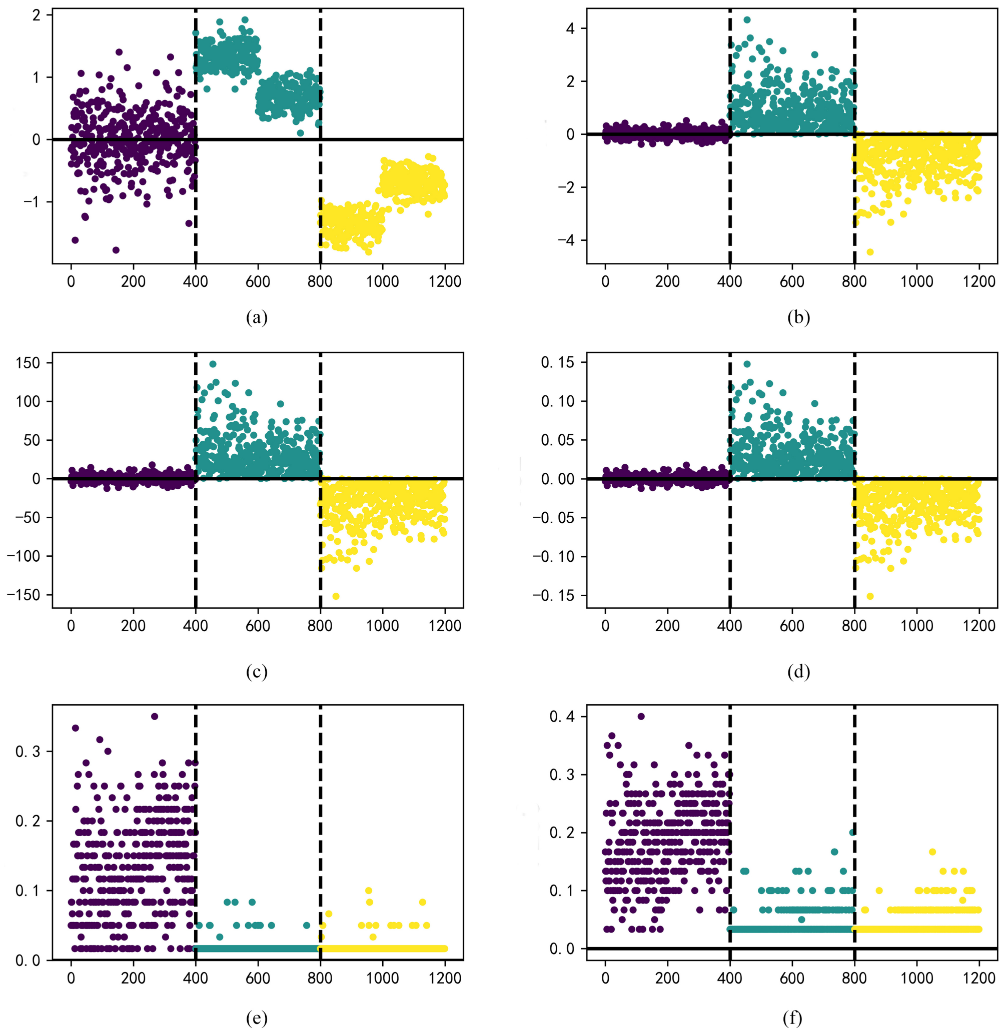

Figure 5. Distribution of the first set of features: (a) skew; (b) Vm; (c) S; (d) ASL; (e) PMLC; and (f) PLSC.

In this way, the intelligent recognition of the control chart pattern is completed.

The value of SVM super parameter in two-layer classifier will affect the accuracy of recognition. FOA is used to optimise the two super parameters C and g of SVM to improve the performance of the classifier. The concrete steps of pattern recognition model based on Sup-FOA-MSVM control chart are shown in Algorithm 2 in Table 2.

3. Results

Six common control chart modes are studied. According to the mathematical equations and parameter ranges shown in Table 3, 400 sets of data samples are generated for each control chart mode using the Monte Carlo simulation method. Among them, 200 groups are used for training, and 200 groups are used for testing. There are total 2,400 groups of simulation data samples, and the length of the control chart window is set to 60.

In Table 3, $y(t)$ represents the observed value at the sampling time point t; μ is the sample mean, and $x(t)$ is the objective random disturbance in the manufacturing process, which obeys normal distribution. d is the slope of the upward trend and downward trend, and b is the step position in the step pattern. $b = 0$ before the step, and $b = 1$ after the step. s is the step amplitude; a is the change amplitude of the period in the periodic mode, and T is the change period of the periodic mode. The mean μ is set to 0. To better simulate the contingency and uncertainty of the random disturbance of different batches of products or process parameters in actual manufacturing, the standard deviation σ is set to be uniformly distributed in the range. The abnormal starting points of the periodic pattern follow a uniform distribution in the range. The starting points of trend pattern and step pattern are evenly distributed in the range. The slope d in the trend pattern is in the set $\{0.05\sigma,\ 0.1\sigma,\ 0.15\sigma,\ 0.2\sigma,\ 0.25\sigma\}$. In the step mode, the step amplitude s is in the range $\{\sigma,\ 1.5\sigma,\ 2\sigma,\ 2.5\sigma\}$. In the periodic pattern, the amplitude a is in the set $\{\sigma,\ 1.5\sigma,\ 2\sigma,\ 2.5\sigma\}$, and the period T is in the range $\{12,\ 14,\ 16\}$. To better simulate the complexity of real sample data changes, all these parameters follow uniform distribution in their ranges.

The parameters of four SVM classifiers in the two-layer classifier are optimised by using FOA. The number of parameters N for FOA optimisation is set as 2, that of iterations as 100, and the population size P as 30.

Table 2
Sup-FOA-MSVM Control Chart

Algorithm 1. Based on Sup-FOA-MSVM recognition model
Input: the original data, the number of parameters N, the number of iterations I, the population size P, the random initial position of the drosophila swarm $(X,\ Y)$, and the random flight direction and distance of individual drosophila. Output: the results of control chart pattern recognition and the classification accuracy of cognition model. Step1: preprocess the original data using the (1) and then save the processed results in the data set (DS). Step2: extraction features of skew, Vm, S, ASL, PMLC and PLSC from DS, then generate the first feature set (FS1). Step3: optimise the two super parameters $C0$ and $g0$ in decision layer 1 by the Fruit fly optimisation algorithm and obtain the optimum parameter of $C0$ and $g0$. Step4: put the optimum parameter of $C0$ and $g0$ into the MSVM, train and test based on the FS1 to obtain the classification result and classification accuracy. Step5: extraction features of Vs, PMLC, PLSC, ASL, ACL and REAE from the classification result to generate the second feature set (FS2). Step6: divide FS_2 into flat mode data set FS21, ascending mode data set FS22 and descending mode data set FS23 based on the prediction tag of classification result. Step7: optimise the parameters of the classifier SVM1, SVM2 and SVM3 in decision level 2 by the fruit fly optimisation algorithm, and obtain the best parameters $C1$, $g1$, $C2$, $g2$ and $C3$, $g3$. Step8: put the FS_{21}, FS_{22}, FS_{23} obtained from the Step6 and the optimum parameter of $C1$, $g1$, $C2$, $g2$, $C3$, $g3$ obtained from the Step7 into the SVM1, SVM2 and SVM3, train and test to obtain the classification result. Step9: calculate the classification accuracy of Sup-FOA-MSVM recognition model.

Table 3
Numerical Equations and Parameter Ranges
of Six Control Chart Modes

Models	Mathematical Equations	Parameters and Ranges
NOR	$y\left(t\right)=\mu+x\left(t\right)$	$\mu=0, 0<\sigma\leq 0.5$
CYC	$y\left(t\right)=\mu+x\left(t\right)+a\times\sin(2\pi t/T)$	a $\epsilon\left\{\sigma, 1.5\sigma, 2\sigma, 2.5\sigma\right\}, T\ \epsilon\left\{12, 14, 16\right\}$
UT	$y\left(t\right)=\mu+x\left(t\right)+d\times t$	d $\epsilon\left\{0.05\sigma, 0.1\sigma, 0.15\sigma, 0.2\sigma, 0.25\sigma\right\}$
US	$y\left(t\right)=\mu+x\left(t\right)-d\times t$	
DT	$y\left(t\right)=\mu+x\left(t\right)+b\times s$	s $\epsilon\left\{\sigma, 1.5\sigma, 2\sigma, 2.5\sigma\right\}, b\ \epsilon\left\{0, 1\right\}$
DS	$y\left(t\right)=\mu+x\left(t\right)-b\times s$	

Table 4
Optimisation Results of FOA-SVM

Classifiers	Optimisation Parameters		Classification Accuracy (%)
	C	g	
MSVM	0.405	0.433	99.8
SVM1	7.694	7.080	94.7
SVM2	3.605	0.334	98.0
SVM3	0.989	0.816	97.7

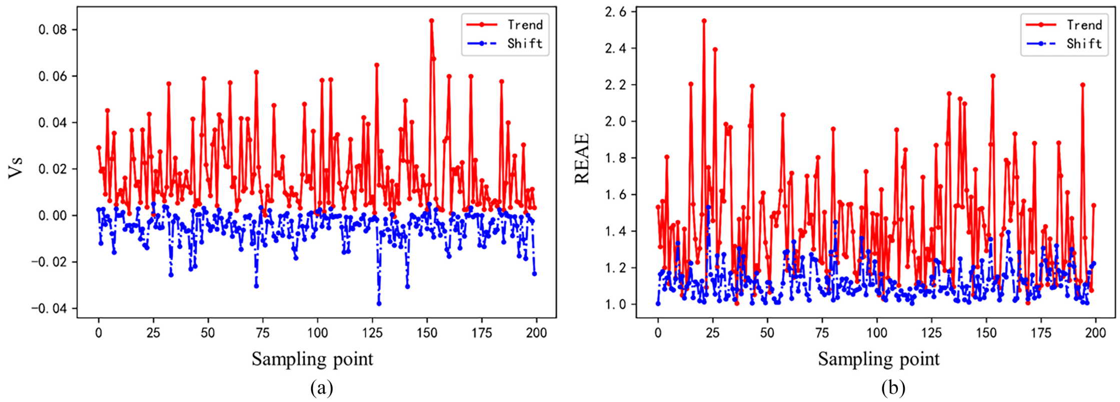

Figure 6. Vs and REAE comparison between trend mode and shift mode.

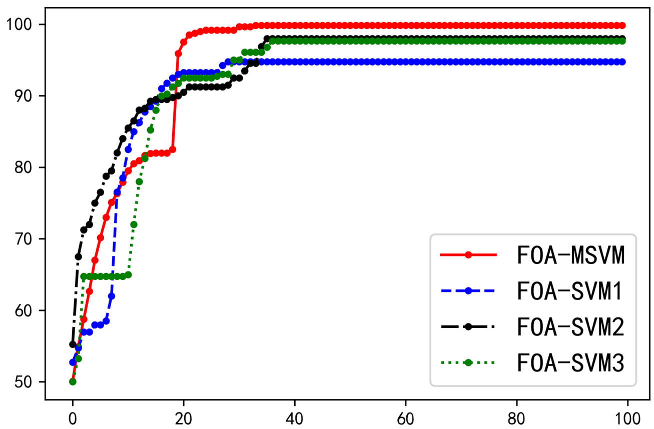

Figure 7. Optimisation curve of FOA-SVM.

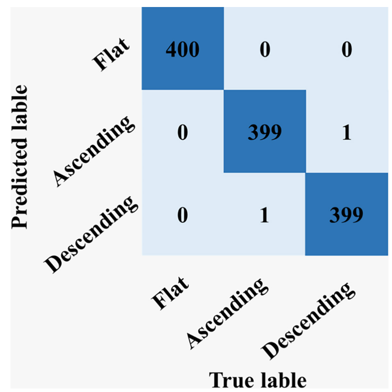

Figure 8. Confusion matrix of FOA-MSVM model for classification of flat, ascending, and descending patterns.

The random initial position range of drosophila swarm is between (0, 1). The super parameters C and g are in the range of (0,100). Figure 7 shows the optimisation curve of 50 fold cross verification of four SVM classifiers in the two-layer classifier. After optimisation, the best parameters C and g and the highest classification accuracy of the four SVM classifiers are shown in Table 4.

To verify the effectiveness of the preprocessing feature extraction based on superposition method, the recognition accuracy of classifier is improved. The original data (60-dimension, SVM1) is used to extract the first set of features (60-dimension, SVM2) and the superposition method is used to preproccess the data from the original data (60-dimension, SVM3), so as to analyse and compare the classification accuracy is shown in Table 5. Multi-classification support vector machine MSVM is used in all methods. Gaussian kernel function is selected for MSVM, and the super parameters C and g are the optimal parameters obtained after fruit fly algorithm optimisation.

The confusion matrix of the FOA-MSVM model (as shown in Fig. 8) has remarkable classification. For the "Flat" category, all 400 samples are correctly predicted, demonstrating the capability of the model in identifying "Flat" mode. In the "Ascending" category, 399 samples are accurately classified, with only 1 misjudged as "Descending"; symmetrically, the "Descending" category generates 399 correct predictions and 1 misclassification as "Ascending." The minimal error between "Ascending" and "Descending" may stem from subtle feature similarities in specific dynamic scenarios. Overall, the model demonstrates high accuracy and can effectively distinguish diverse pattern categories. Negligible misclassifications also provide insights for future research-further refining feature extraction to address subtle inter-category ambiguities could enhance the robustness of the model.

Considering the uncertainty of the position of the standard deviation σ of the sample data and the abnormal starting point t_0, the proposed model and other control models are compared, as shown in Table 6. Model 1 is based on the original data support vector machine; model 2 is based on the original data to extract statistical features and shape features support vector machine. As shown in Table 6, when the standard deviation σ and the abnormal starting position t_0 of the sample data are unchanged, the recognition accuracy of the three models is high. When the standard deviation σ and the abnormal starting position t_0 of the sample data change, the recognition accuracy of the first two models decreases significantly. However, the proposed model still has a high recognition accuracy of 96.5%. The model has advantages to address pattern recognition in control charts, where the position of standard deviation σ and the starting point t_0 of the abnormal pattern constantly changes, which is conducive to the subsequent anomaly analysis and troubleshooting.

Table 5
Methods Comparison of Recognition Performance of Different MSVM

Models	Correct Classify Number Based on Different Model			Classification Accuracy (%)
	Flat Mode	Ascending Mode	Descending Mode	
FOA-SVM1	397	357	351	92.1
FOA-SVM2	390	384	382	96.3
FOA-SVM3	381	380	381	95.2
FOA-MSVM	400	399	399	99.8

Table 6
Recognition Comparison of Different Models

Models		Correct Classify Number Based on Different Models						Classification Accuracy (%)
		NOR	CYC	UT	US	DT	DS	
Fixed parameters σ and t_0	Model1	196	195	179	191	187	193	94.8
	Model2	195	196	191	194	191	199	97.2
	Model of our work	200	197	197	199	196	196	98.7
Changed parameters σ and t_0	Model1	190	109	185	170	188	178	83.3
	Model2	175	170	193	172	190	176	89.7
	Model of our work	188	190	195	196	196	193	96.5

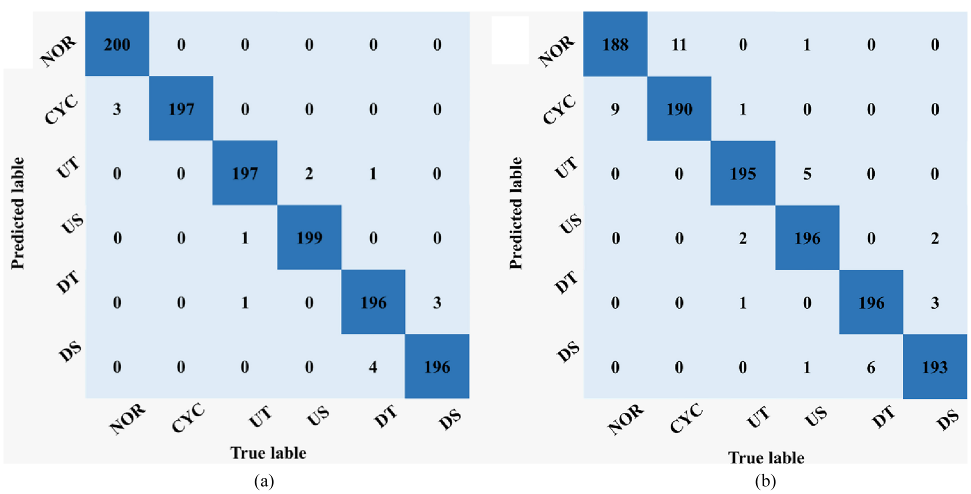

Figure 9. Confusion matrices of pattern classification model.

Compared with single-tier SVM, the two-tier architecture can reduce computational complexity. For example, the proposed model needs 128 s for the training on 2,400 samples, whereas a single-tier SVM requires 182 s. A 30% reduction in computational time indicates that the selected model has high computational efficiency.

The confusion matrices are shown in Fig. 9(a) and 9(b). In the ideal parameter scenario of Fig. 9(a), the model can achieve 100% correct classification for the NOR pattern (200/200), and show dominant in other patterns (e.g., 197 correct classifications for CYC). Only a few samples are misjudged, reflecting the precise identification capability of the model under standard parameter settings. Figure 9(b) shows the classification under dynamic parameter variations. Although certain misclassifications occur in NOR and CYC, the model still maintains a high correct classification rate, such as 196 correct classifications for the U.S. pattern. Combining the advantages of the model, its hierarchical classifier architecture and optimisation strategy (*e.g.,* FOA-based parameter optimisation) play a key role: First, the hierarchical design achieve progressive feature discrimination, improving the ability to distinguish complex patterns. Second, the optimisation mechanism can improve the adaptability to parameter fluctuations. Compared with other comparative models, this model can better maintain classification stability in dynamic industrial scenarios and effectively address recognition challenges caused by parameter variations in statistical process control. This highlights its advantages in handling complex and dynamic patterns for practical engineering applications.

4. Conclusion

A Sup-FOA-MSVM control chart pattern recognition model is proposed. The data preprocessing method based on sample information superposition is used to highlight the significance features of the original data and effectively reduce the dimensions of original data. Then a two-layer support vector machine classifier is designed based on feature analysis. The fruit fly optimisation algorithm is used to optimise the hyperparameter of the classifier model, which can improve the recognition accuracy of the model. Through simulation comparison, the proposed model can effectively solve the problem of pattern recognition of control chart. Especially in solving the problem of standard deviation of sample data and uncertainty of abnormal starting point position, the proposed model shows high recognition accuracy and practical application value. Future research would explore deep learning integration for hierarchical feature extraction and real-world validation on unbalanced industrial datasets to address domain-specific challenges.

Acknowledgment

We thanks Ministry of Education industry-university cooperative education project, grant number 220906470201746, University Research Project of Anhui Province, grant number 2023AH052275, Quality engineering project of Bozhou university, grant number 2023XJXM001, and The project of Bozhou University BYKQ202301.

References

[1] Z. Q. Wang and Z. He, Control charts for attribute data: Review and perspectives, *Journal of Systems Engineering, 39*(3), 2024, 344–359. DOI: https://doi.org/10.13383/j.cnki.jse.2024.03.003.

[2] W. Zhong and L. Liu, High-dimensional nonparametric EWMA control chart based on empirical likelihood ratio test, *Journal of Systems Science and Mathematical Sciences, 44*(3), 2024, 862–878. DOI: https://doi.org/10.12341/j.ssms.122875.2024.03.010.

[3] X. M. Zhang and Y. Fu, Adaptive model predictive control based on compensation control, *Theory & Applications, 11*, 2024, 1–8. DOI: http://kns.cnki.net/kcms/detail/44.1240.TP.20231214.0848.042.html.

[4] M. D. Zhang, Automatic monitoring system for tobacco silk processing equipment based on statistical process contro (SPC), *Automation & Instrumentation, 39*(4), 2024, 47–51. DOI: https://doi.org/10.19557/j.cnki.1001-9944.2024.04.010.

[5] W. Ning, Z. Guo, S. Tian, and K. Li, Control chart quality abnormal pattern recognition based on hybrid feature t-SNE fusion, *Systems Engineering Theory & Practice, 44*(7), 2024, 2381–2393. DOI: https://doi.org/10.12011/SETP2023-157.

[6] H. A. O. Lanyu, Z. Di, L. I. Yanting, and P. A. Ershun, Multivariate coupled statistics monitoring of wafer manufacturing overlay errors based on copula, *China Mechanical Engineering, 34*(3), 2023, 369–377. DOI: https://doi.org/10.3969/j.issn.1004-132X.2023.03.014.

[7] W. Hachicha and A. Ghorbel, A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme, *Computers & Industrial Engineering, 63*(1), 2012, 204–222. DOI: https://doi.org/10.1016/j.cie.2012.03.002.

[8] H. Y. Wang and Y. J. Zhuo, Pattern recognition method of PCA-based statistical process control chart, *Statistics & Decision, 36*(24), 2020, 20–24. DOI: https://doi.org/10.13546/j.cnki.tjyjc.2020.24.004.

[9] A. Addeh, A. Khormali, and N. A. Golilarz, Control chart pattern recognition using RBF neural network with new training algorithm and practical features, *ISA Transactions, 79*, 2018, 202–216. DOI: https://doi.org/10.1016/j.isatra.2018.04.020.

[10] T. T. El-Midany, M. A. El-Baz, and M. S. Abd-Elwahed, A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks, *Expert Systems with Applications, 37*(2), 2010, 1035–1042. DOI: https://doi.org/10.1016/j.eswa.2009.05.092.

[11] J. Xu, H. Lv, Z. Zhuang, Z. Lu, D. Zou, and W. Qin, Control chart pattern recognition method based on improved one-dimensional convolutional neural network, *IFAC-PapersOnLine, 52*(13), 2019, 1537–1542. DOI: https://doi.org/10.1016/j.ifacol.2019.11.418.

[12] T. Zan, Z. Liu, H. Wang, M. Wang, and X. Gao, Control chart pattern recognition using the convolutional neural network, *Journal of Intelligent Manufacturing, 31*, 2020, 703–716. DOI: https://doi.org/10.1007/s10845-019-01473-0.

[13] S. K. Gauri and S. Chakraborty, Recognition of control chart patterns using improved selection of features, *Computers & Industrial Engineering, 56*(4), 2009, 1577–1588. DOI: https://doi.org/10.1016/j.cie.2008.10.006.

[14] A. Anani, S. O. Adewuyi, N. Risso, and W. Nyaaba, Advancements in machine learning techniques for coal and gas outburst prediction in underground mines, *International Journal of Coal Geology, 285*, 2024, 104471. DOI: https://doi.org/10.1016/j.coal.2024.104471.

[15] S. Dong, L. Yu, W. A. Zhang, and B. Chen, Robust extended recursive least squares identification algorithm for Hammerstein systems with dynamic disturbances, *Digital Signal Processing, 101*, 2020, 102716. DOI: https://doi.org/10.1016/j.dsp.2020.102716.

[16] X. Zhou, P. Jiang, X. and Wang, Recognition of control chart patterns using fuzzy SVM with a hybrid kernel function, *Journal of Intelligent Manufacturing, 29*, 2018, 51–67. DOI: https://doi.org/10.1007/s10845-015-1089-6.

[17] S.F. Ahmed, M. S. B. Alam, M. Hassan, M. R. Rozbu, T. Ishtiak, N. Rafa, M. Mofijur, A. B. M. Shawkat Ali, and A. H. Gandomi, Deep learning modelling techniques: current progress, applications, advantages, and challenges, *Artificial Intelligence Review, 56*, 2023, 13521–13617. DOI: https://doi.org/10.1007/s10462-023-10466-8.

[18] E. Kara, G. Zhang, J. J. Williams, G. Ferrandez-Quinto, L. J. Rhoden, M. Kim, J. N. Kutz, and A. Rahman, Deep learning based object tracking in walking droplet and granular intruder experiments, *Journal of Real-Time Image Processing, 20*, 2023, 86. DOI: https://doi.org/10.1007/s11554-023-01341-4.

[19] D. Hu, X. F. Li, X. F. Jiang, R. L. Yang, Y. Liu, and M. Dong, Learning fusion diagnosis method of partial discharge in oil-paper insulation based on multi-spectral information, *Advanced Technology of Electrical Engineering and Energy, 43*(10), 2024, 85–92. DOI: https://doi.org/10. 12067/ATEEE2305020.

[20] Y. Yang, H. Chen, S. Li, A. A. Heidari, and M. Wang, Orthogonal learning harmonising mutation-based fruit fly-inspired optimizers, *Applied Mathematical Modelling, 86*, 2020, 368–383. https://doi.org/10.1016/j.apm.2020.05.019.

[21] T. Liu, J. Feng, D. Fu, and W. Chen, Application of the optimised YOLOv3 algorithm in substation power equipment defect identification and detection, *Mechatronic Systems and Control, 51*(4), 2023, 235–243. DOI: 10.2316/J.2023.201-0386.

[22] S. Song, Low energy office building design based on non-dominated sorting genetic algorithm 2 and extreme gradient boosting-artificial neural network, *Mechatronic Systems and Control, 52*(3), 2024, 194–202. DOI: 10.2316/J.2024.201-0468.

Biographies

Juanjuan Qin received the master's degree in electrical engineering from Anhui University of Science and Technology, Huainan, Anhui, China, in 2012. She works with the Department of Electronic and Information Engineering, Bozhou University as a Senior Experimental Engineer. Her research focuses on intelligent control algorithms, pattern recognition, and the safety of industrial control systems.

Baokang Zhang received the B.E. degree in electrical engineering and automation from Yangzhou University, China, in 2019. He is currently pursuing the Ph.D. degree in electronic information from Zhejiang University of Technology, Hangzhou, China. His research focuses on industrial fault diagnosis, privacy-preserving data analysis, and intelligent anomaly detection systems.

Qi Wu received the Ph.D. degree in control theory and control engineering from the Zhejiang University of Technology, Hangzhou, China, in 2021. He is currently a Postdoctoral Researcher with the College of Information Engineering, Zhejiang University of Technology. His research interests include networked motion control systems, industrial data analysis, privacy analysis, and protection with applications to industrial control system and power systems.

Ning Sun received the B.S. degree in electrical engineering and automation from Changchun University, Changchun, China, in 2020; and the M.S. degree in electrical engineering and automation from Henan Polytechnic University, Jiaozuo, China, in 2023. He is currently working as a Lecturer with the School of Electronic Information Engineering, Anhui Bozhou University. His current research interests include power converter control and microgrid stability analysis.

Maoqi Cao received the Ph.D. degree from the University of Science and Technology of China, in 2015. He currently works as a Professor of electronic and information engineering with Bozhou University. His current research interests include functional materials and dynamics.