Damage Identification of Prestressed Concrete Components Based on Machine Learning Optimization Algorithm and Piezoelectric Wave Measurement

Hongtao Zhu[1], Shuyun Guo[2*]
[1]School of Civil Engineering, Xinyang College, Xinyang 464000, China
[2]School of Foreign Languages, Gushi Vocational Education Center, Xinyang 464000, China
E-mail: zhuhongtao2008hi@163.com, 15713761803@163.com
*Corresponding author

Prestressed concrete components have high resistance to cracks and stiffness, yet prone to damage leading to accidents under adverse environments and extreme loads. The study uses machine learning algorithms to construct an intelligent concrete damage recognition model aimed at accurately assessing its health status. The piezoelectric wave measurement method is used to collect small wave signals from concrete. The improved back propagation network is used to identify concrete damage characteristics in the signals, and the support vector machine is taken to correct the identification results. The genetic algorithm is used to optimize the back propagation neural network, obtaining the optimal threshold and weight of the back propagation neural network to improve its robustness and feature extraction ability in noisy environments. According to the results, the model constructed by integrating two classification algorithms has a mean square error of 7.962×10^{-4}, a coefficient of determination of 0.9756, and an F1 score of 0.9836 in the damage location recognition results. In the identification results of the degree of damage, the mean square error of the research model was 6.548×10^{-2}, the coefficient of determination was 0.9531, and the F1 score was 0.9925, respectively. In the environment with introduced noise, the recognition accuracy of the research model was 93.7%. The results indicate that the research method has higher accuracy and robustness in damage identification compared with other models, which can be used for concrete damage detection in large buildings or long-term high load buildings.

Povzetek: A hybrid GA-BP-SVM model enhances damage identification in prestressed concrete using piezoelectric wave measurement, achieving high accuracy (F1: 0.9925) and robustness (93.7% under noise), improving structural health monitoring.

1 Introduction

Concrete materials have abundant raw materials, low prices, simple production processes, durability, and strong plasticity, which are widely used in infrastructure construction and building construction. With the large-scale and diversified development of civil engineering, concrete may be subjected to strong external pressure or long-term corrosion, causing damage and leading to structural safety in buildings [1]. However, the early characteristics of concrete damage are not obvious. It is difficult to fully cover them through regular manual inspections, which cannot provide accurate results for assessing its health status. The piezoelectric wave measurement method uses piezoelectric materials to generate and detect waves. The instrument is small, easy to use, and not affected by the environment. It can be used for monitoring and warning, improving structural safety. The piezoelectric wave measurement method combined with machine learning algorithms can adaptively detect concrete damage, with high efficiency and accuracy. Back Propagation (BP) network can handle complex nonlinear problems and automatically extract patterns between data. It has high adaptability and self-learning ability, which is a commonly used feature classification algorithm. Genetic Algorithm (GA) simulates the natural selection process and has an automatic elimination mechanism. Through selection, hybridization, and mutation, the population evolves gradually towards the optimal solution. It is a meta-heuristic algorithm used for optimization. Support Vector Machine (SVM) is an extensively applied supervised learning algorithm. SVM has high accuracy in classification and regression tasks, especially in high-dimensional spaces, which has high generalization ability. Many scholars have conducted relevant research on piezoelectric wave measurement methods, GA, BP, and SVM algorithms.

In response to the inaccurate measurement of explosion shock wave pressure under the instantaneous high temperature effect in the explosion field, Shi et al. used piezoelectric wave sensors to analyze the effects of environmental temperature and transient temperature. A theoretical analysis method for transient temperature was proposed. A transient temperature control strategy was designed by coating 0.5 mm thick lubricating silicone oil on the sensor membrane and 0.2 mm thick glass fiber cloth on the sensor side. The accuracy of the explosion shock wave pressure was improved to 97.8% [2]. To accurately predict whether the surface roughness of

precision manufactured aluminum alloys meets the requirements, Bai et al. designed a BP prediction method. The results showed that the predictive progress of the model was improved by 7.8%. BP recognized small features in high-precision manufacturing and solved the insufficient accuracy [3]. In response to the inaccurate prediction of the self-diffusion coefficient in pure liquids, Zeng et al. used the BP to establish a nonlinear method that predicted the self-diffusion coefficient of pure liquids at normal pressure. The R2, AARD, and RMSE for predicting the self-diffusion coefficient were 0.9940, 7.09%, and 0.1106, respectively [4]. Jiang et al. used an optimized GA to plan the optimal layout of the clothing production line for the uneven work intensity among employees. The results showed that the balance rate of the production line increased from 70.5% to 97.05%, and the production cycle was shortened by 32.80%, verifying the efficiency improvement performance [5]. To solve the difficulty in early warning of nonlinear macroeconomic systems, Yin et al. proposed a CNN economic early warning system based on the IGA-BP. The correlation coefficient was 0.89, and the delay number was 0. The economic warning system based on IGA-BP algorithm was effective. The BP optimized by GA could improve its feature classification ability [6]. To optimize the ability to prevent and control large-scale crop diseases, Gangadevi et al. proposed an improved SVM plant disease and pest identification model. The recognition accuracy of the research model reached 91.1%, which was superior to other models. The powerful feature classification ability of SVM could solve the crop disease prevention and control [7]. Dong et al. proposed a runoff prediction model based on SOA-SVM to address the difficult runoff prediction caused by its nonlinear and non-stationary characteristics. The results indicated that the average error and RMSE indicators of the research model were superior to other models [8]. Zhang et al. designed a method for automatically extracting damage characteristics from point cloud data to address the inaccurate damage detection in reinforced concrete structures. The results indicated that the research method was an effective approach for post disaster impact assessment and large-scale building damage detection [9].
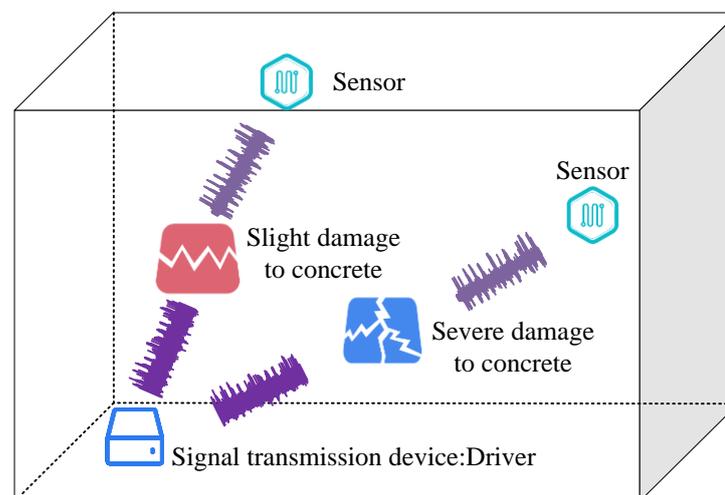
Figure 1: The principle of piezoelectric wave measurement

The above research indicates that existing detection methods have insufficient ability to detect small cracks and other damages in concrete structures in the early stages. Although existing literature has made significant progress in the application of machine learning and optimization algorithms, relatively little research has been done in the field of damage identification of prestressed concrete members, especially in the combination of piezoelectric wave measurement and GABP-SVM model. Existing methods often do not perform well when dealing with noise and uncertainty in complex environments. The piezoelectric wave measurement method can identify extremely small wave changes, providing a data basis for damage detection of concrete components. GA can be applied to optimize the BP for extracting damage features from piezoelectric wave measurement data. Therefore, introducing SVM algorithm for feature classification can improve the accuracy of damage detection in prestressed concrete components. The research aims to combine piezoelectric wave measurement method with GA, BP, and SVM algorithms to construct a new intelligent damage identification model for prestressed concrete components, monitoring the health status of concrete.

The research is structured from three sections. The first introduces related algorithms and mechanisms in the concrete damage recognition model based on GABP-SVM. The second section tests the model. The third section summarizes the research results. The latest research and comparison of research results in this field are shown in Table 1.

Table 1: Comparison chart of SOTA and research achievements in this field

Algorithms	Year	Researcher	Method	MSE	R2	F1	Prediction accuracy	SOTA lacks
Optimize BP	2023	Bai et al[3]	BP	/	/	/	/	Traditional BP may fall into local optima
	2022	Zeng et al[4]	BP-ANN	/	0.9940	/	/	Not tested for robustness in complex environments
	2022	Yin et al[6]	IGA-BP	/	/	/	/	Robustness unverified
	/	This study	Improved GA-BP	/	0.9235	0.00724	/	/
Research model	2024	Gangadevi et al[7]	FOA-SVM	/	/	0.945	91.1%	Unoptimized FOA may obtain non optimal solutions
	2023	Dong et al[8]	SOA-SVM	/	/	/	/	Robustness unverified
	2022	Zhang et al[9]	Point cloud data	/	/	/	/	Low efficiency and unstable accuracy
	/	This study	GA-BP-SVM	6.548×10^{-2}	0.9531	0.9836	93.7%	/

2 Methods and materials

The study first introduces the principle of electromagnetic wave dynamic measurement method. Then, the GA is taken to optimize BP to improve the BP algorithm to identify concrete damage characteristics from piezoelectric wave measurement data. Finally, SVM is used to further classify and modify the output of GABP, improving the accuracy of damage detection in prestressed concrete components.

2.1 Design of concrete damage feature classification algorithm based on GABP

Under earthquake, high load and other conditions, cracks and other damages may occur on the surface or inside of prestressed concrete components. When cracks are small or damage occurs inside concrete components, they are usually difficult to detect. If not repaired timely, it may lead to major accidents such as structural collapse. Piezoelectric smart materials have positive/negative piezoelectric effects. The positive piezoelectric effect is manifested by the internal polarization when subjected to external forces, releasing charges proportional to the pressure. The negative pressure electric effect is manifested in the external electric field, where materials convert electrical energy into mechanical energy and undergo deformation [10-11]. Therefore, piezoelectric smart materials can be used as signal sensors, which are taken as signal transmission device in concrete structure damage detection. The principle of using piezoelectric intelligent materials for piezoelectric wave measurement is shown in Figure 1.

In Figure 1, the signal transmission device deforms under the action of an electric field to generate sinusoidal stress waves, which propagate inside the concrete. During the propagation process, the signal will gradually decay due to energy loss caused by friction, scattering, and crack absorption between the medium and particles.

Therefore, there is the energy conservation $E_I = E_H + E_R + E_T$. E_I is the energy of stress waves, which is the initial stress wave energy. E_H represents energy loss, which refers to the energy lost during the propagation of stress waves in concrete due to friction, scattering, and absorption. E_R is the energy of the transmitted wave received by the sensor after passing through the concrete. E_T is the reflected wave energy, which refers to the energy reflected back by stress waves when they encounter damage or interfaces inside concrete. Therefore, by collecting signals through sensors, concrete damage can be monitored. There is much interference noise in the signals collected by sensors. Many options for dimensionality reduction can preserve the main information. Singular value decomposition can select the k largest singular values to explain most of the variance in the signal, preserving the main information. Therefore, the study uses singular value decomposition to reduce the dimensionality of the signals, set the threshold of cumulative variance contribution to 90%, and introduces the BP algorithm to identify the characteristics of concrete damage information in the signals. The BP is displayed in Figure 2.

In Figure 2, during the forward propagation process, initialization is first performed, and then the sample data enters the hidden layer for calculation using a transfer function. The result is then fed to the output layer, where the error and calculation result are calculated. Finally, the termination condition is reached and the prediction result is output. During the iteration process, if the termination condition is not satisfied, BP is performed, and updated the obtained training error to all neurons. Each neuron adjusts the weights and thresholds of the entire network on the basis of the training error. If the training frequency reaches the set condition or the error reaches the

minimum, the neural network ends. Otherwise, it enters the hidden layer to continue calculation until the condition is met. The forward propagation is shown in equation (1) [12].

$$\begin{cases} r_q = f_1(H_{R\varphi}, E), \varphi = 1, 2, ..., l \\ o_s = f_2(H_{\varphi s}, A), s = 1, 2..., N \end{cases} \quad (1)$$

In equation (1), r_q is the output of the hidden layer. o_s is the output of the output layer. E is the input sample, A is the output of the hidden layer. N is the sample size. $H_{R\varphi}$ and $H_{\varphi s}$ represent the weight matrices of the hidden layer and the output layer. f_1 is the activation function of the hidden layer. f_2 is the activation function of the output layer. φ is the neuron index of the hidden layer. s is the neuron index of the output layer. The training error is shown in equation (2).

$$J = \frac{1}{2} \sum_{\varphi=1}^{N} \sum_{s}^{m} (i_{ns} - o_{ns})^2 \quad (2)$$

In equation (2), m satisfies the output layer node. n satisfies the input layer node. i_{ns} satisfies the expected output of the network. J represents the training error. $\frac{1}{2}$ is the coefficient used in error calculation. i_{ns} is the expected output value of the network for the i-th output node and the s-th sample. o_{ns} is the output value of the n-th neuron in the output layer for the s-th sample. The updated network weights are shown in equation (3).

$$\begin{cases} \varpi_{9s}^{\upsilon+1} = \varpi_{9s}^{\upsilon} + \psi \sigma_k^K r_q, s = 1, 2, ..., S; d = 1, 2, ..., L \\ \varpi_{g9}^{\upsilon+1} = \varpi_{g9}^{\upsilon} + \psi \sigma_d^L E_g, l = 1, 2, ... L; g = 1, 2, ..., G \end{cases} \quad (3)$$

(a) Exchange variation   (b) Insertion variation   (c) Flashback variation
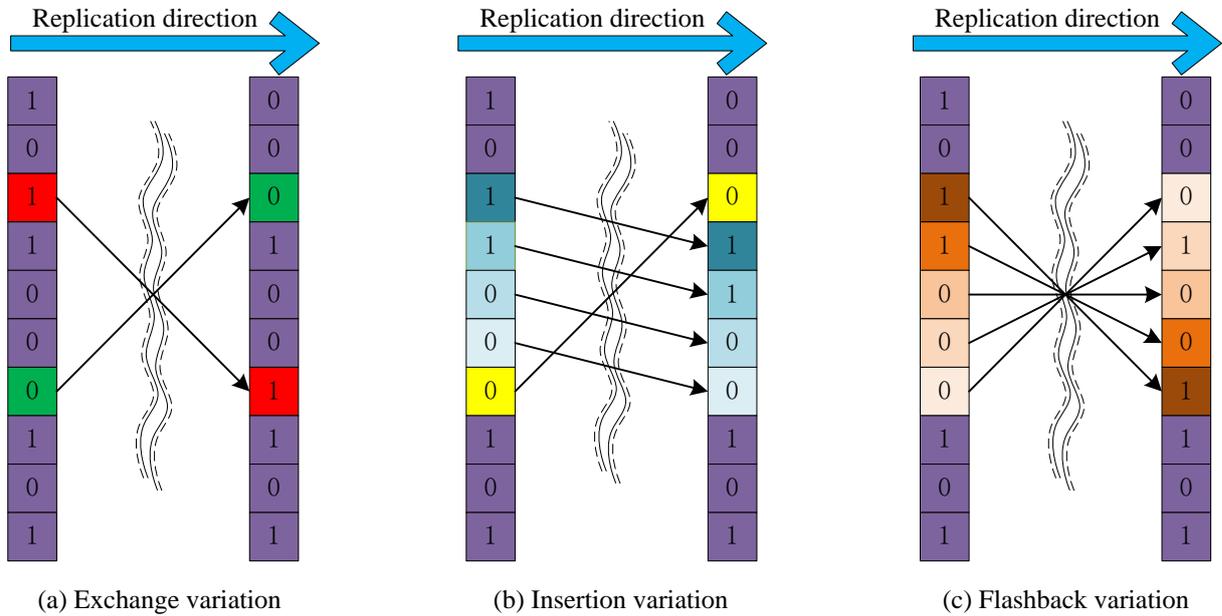
Figure 3: Genetic algorithm variation operation

In equation (3), $\varpi_{9s}^{\upsilon+1}$ represents the weight matrix of the $\upsilon+1$-th iteration from the hidden layer 9 to the output layer s. ϖ_{9s}^{υ} is the weight matrix before the update. S is the maximum number of output layers. σ_k^K is the specific layer weight update for the k-th iteration. K is the total number of iterations. ϖ_{g9}^{υ} is the weight matrix from the input layer g to the output layer 9 during the υ-th iteration. $\varpi_{g9}^{\upsilon+1}$ is the updated weight matrix. G is the total number of input layers. σ_d^L is the weight update of the output layer for the l-th iteration. L is the maximum number of iterations. ψ satisfies the learning rate. υ is the maximum number of iterations. r_q

is the output of the hidden layer in the q-th iteration. E_g is the input sample in the g-th iteration. Due to the BP over-fitting in noisy environments, there are numerous heuristic algorithms that can be used to optimize BP, among which GA exhibits high stability and robustness when dealing with complex problems, and can better cope with noise and uncertainty. Therefore, the study introduces GA to optimize BP. The GA evaluates the suitability of individuals based on the fitness function and the gap between them and the optimal target. Genetic methods such as selection, crossover, and mutation are performed to simulate natural selection of better individuals. Different fields will use different fitness functions. Generally, the fitness function of GA is transformed from the objective

function to the fitness function, as displayed in equation (4).

$$f(x) = \begin{cases} g(x) \\ -g(x) \end{cases} \quad (4)$$

In equation (4), when solving the minimization problem, the fitness function $f(x) = -g(x)$ is used.

When solving the maximization problem, the fitness function is $g(x)$. When the fitness function is not directly transformed from the objective function, the fitness function to solve the minimization problem is shown in equation (5).
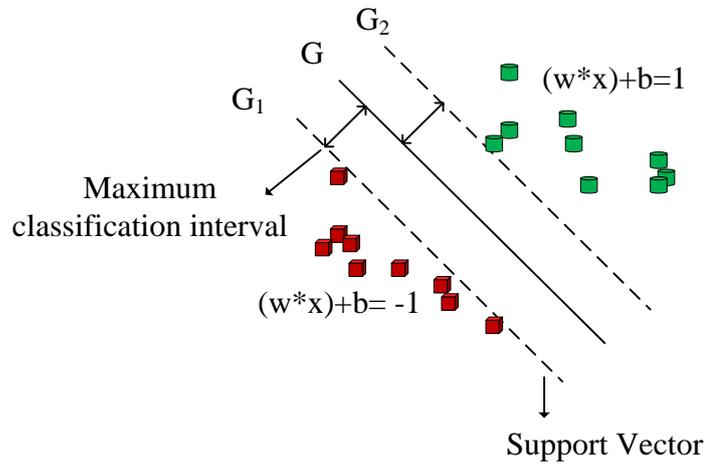
Figure 4: Optimal classification hyperplane for SVM
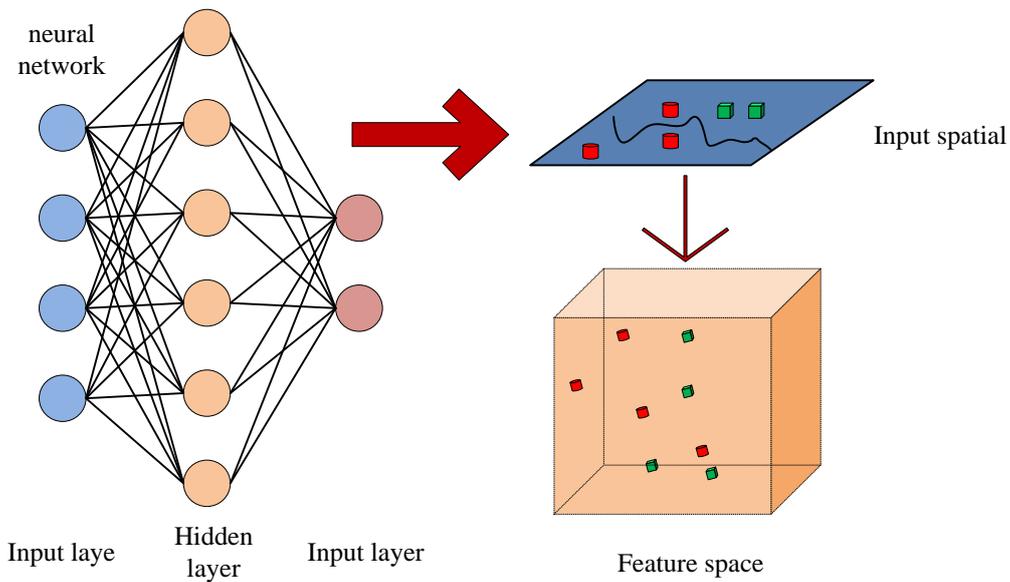
Figure 5: SVM network topology diagram

$$f(x) = \begin{cases} C_{max} - g(x), & g(x) < C_{max} \\ 0, & else \\ f(x) = \dfrac{1}{1 + C + g(x)}, & C \geq 0, \quad C + g(x) \geq 0 \end{cases} \quad (5)$$

$$f(x) = \begin{cases} g(x) - C_{min}, & g(x) < C_{max} \\ 0, & else \\ \dfrac{1}{1 + c - g(x)}, & C \geq 0, \quad c - g(x) \geq 0 \end{cases} \quad (6)$$

In equation (5), C_{max} is the maximum fitness value. C is the fitness value. When solving the maximization problem, the fitness function is shown in equation (6).

In equation (6), C_{min} is the minimum fitness value. c is the fitness threshold. The core of GA operators is crossover and mutation, which bring the entire selection

process closer to the optimal solution. GA randomly generates new offspring in the population through crossover operations. When designing the crossover probability, excellent individuals should avoid cross pairing. This can avoid the disappearance of excellent individuals and ensure that new individuals are close to the optimal solution [13]. Similar to natural genetic rules, when there is only cross inheritance, the diversity of the population is insufficient, and individuals who are too similar can easily slow down or even stop the population evolution speed. Therefore, GA includes mutation mechanisms, which increase population diversity through mutation. The genetic mutation operation is displayed in Figure 3.
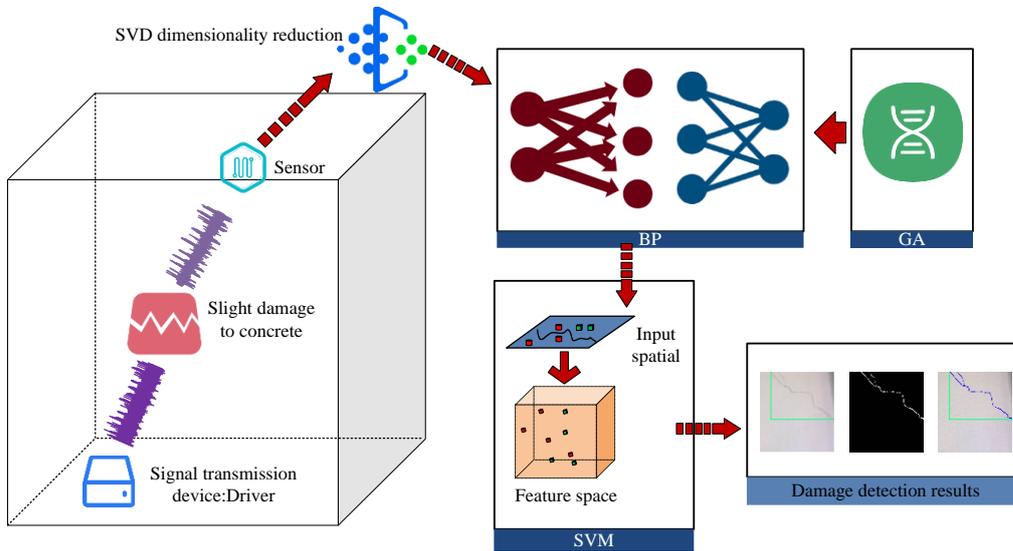
Figure 6: Damage identification model for prestressed concrete components based on GABP-SVM

(a)Prediction results of GABP

(b)Prediction results of BP
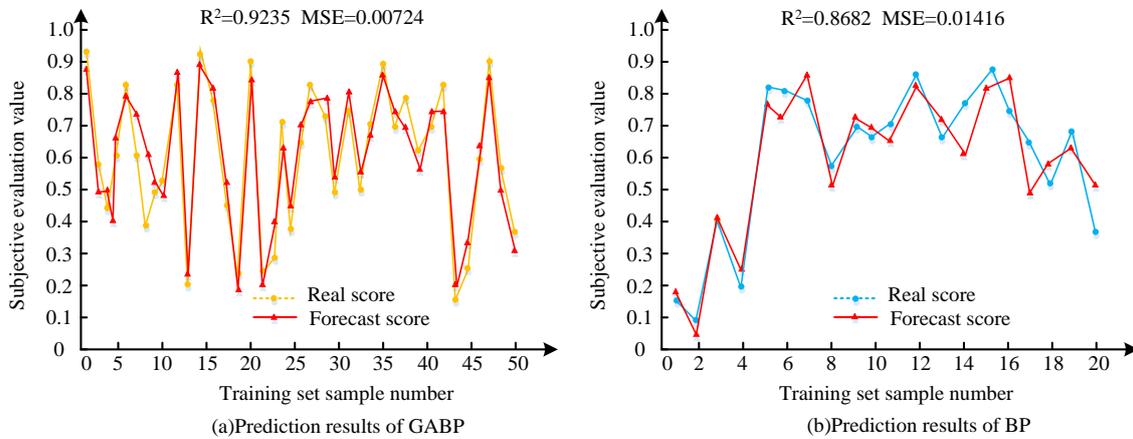
Figure 7: Comparative experimental results of BP and GABP

Figure 3 displays the mutation operation of binary encoding. Figure 3 (a), Figure 3 (b), and Figure 3 (c) depict the exchange, insertion, and flashback during the replication process. Genetic variation produces new individuals, which determines the local search of GA and can bring diversity to the population. These three genetic operators of GA simulate the recombination and mutation process of sexual reproduction, and select better individuals through fitness scores to evolve the entire population towards the optimal solution. The BP optimized by GA is used to map the space vector from n-dimensional to m-dimensional, which can reduce the blindness in the weight adjustment process. The optimized weight correction is shown in equation (7) [14].

$$\begin{cases} w_{jk}(\alpha+1) = w_{jk}(\alpha) + \beta h_k E_r \\ w_{ij}(\alpha+1) = w_{ij}(\alpha) + \gamma i_k E_p \end{cases} \qquad (7)$$

In equation (7), γ and β are learning factors, used to control the magnitude of weight adjustment. $w_{ij}(\alpha)$ and

$w_{jk}(\alpha)$ satisfy the weights of each neuron in the hidden and output layers after α iterations. r is the sample size. E_p is the sample error, used to guide the

adjustment of weights. $w_{ij}(\alpha+1)$ and $w_{jk}(\alpha+1)$ are the weights of each neuron in the hidden layer and output layer after $\alpha+1$ iterations. h_k and i_k are the gradient terms for weight updates, used to minimize the error of the network. The new fitness function is shown in equation (8).

$$F = k\left(\sum_{p=1}^{m} asb\left(o_p - t_p\right)\right) \qquad (8)$$

In equation (8), k is the coefficient. m is the number of output layer nodes. $o_p - t_p$ is the output error of the network. During the optimization process, floating-point encoding is used to encode the basic solution space. After encoding is completed, the population of the GA is initialized. The population M is shown in equation (9).

$$M = m \times n + m \times q + q + m \qquad (9)$$

In equation (9), n is the input layer node size in BP. q is the hidden layer node size in BP. The probability p_i of an individual being selected in GA is shown in equation (10).

$$p_i = \frac{f_i}{\sum\limits_{j=1}^{M} f_j} \qquad (10)$$

In equation (10), $f_i = k / F_i$. F_i represents the fitness value of node i, calculated by equation (8). After determining the individual population F, individuals are decoded and corresponding network connection thresholds and weights are generated. Genetic operations are performed on individuals with lower F-values until the maximum iteration is satisfied to obtain the optimal threshold and weights for the BP network, ultimately optimizing the BP network.

(a) Comparison test results of unimodal function f_1

(b) Comparison test results of multimodal function mf_1
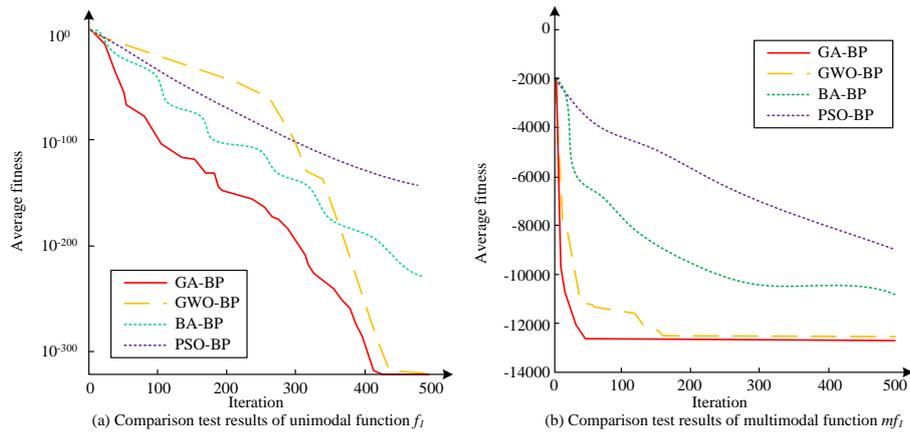
Figure 8: Comparison of optimization algorithms and experimental results

(a) Prediction accuracy of SVM

(b) Prediction accuracy of RF

(c) Prediction accuracy of DT
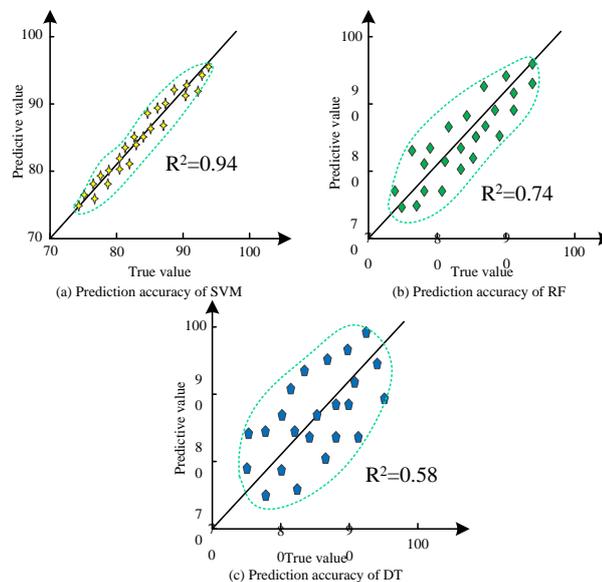
Figure 9: Comparison test results of SVM algorithm and classification algorithm

2.2 Construction of damage identification model for prestressed concrete components based on GABP-SVM

In practical applications, prestressed concrete components may have small cracks in the concrete due to factors such as insufficient compaction, natural shrinkage, and uncleared wood chips, which reduces the GABP feature recognition accuracy [15]. SVM is a supervised learning algorithm widely used in classification and regression analysis. The study introduces SVM to further classify GABP recognition results, aiming to improve the accuracy of concrete component damage detection. In classification problems, SVM attempts to find a hyperplane to maximize the boundary between two categories. The optimal classification hyperplane for SVM is shown in Figure 4.

In Figure 4, in two-dimensional space, the classification hyperplane can be imagined as a line that separates two categories. In a higher dimensional space, the optimal classification hyperplane becomes a hyperplane. SVM can train a classifier for concrete damage feature classification by collecting and analyzing various concrete damage data. The optimal classification plane and lines G, G1, and G2 in SVM algorithm are displayed in equation (11).

$$f(x) = w^T x + b$$

$$(11)$$

In equation (11), w satisfies the normal vector. b satisfies the bias amount. Among all the classified samples, (x_1, y_1), (x_2, y_2), ... and (x_i, y_i) need to satisfy equation (12).

$$y_i(wx_i + b) \geq 1 \qquad (12)$$

In equation (12), different w and b can determine different position planes. The optimal classification plane calculation is shown in equation (13).

$$\begin{cases} \max \dfrac{1}{\|w\|} = \min \dfrac{1}{2}\|w\|^2 \\ s.t, y_i(wx_i + b) \geq 1, i = 1, 2, 3...n \end{cases} \qquad (13)$$

In equation (13), $\min \dfrac{1}{2}\|w\|^2$ represents the minimum confidence range. Equation (13) can be converted into equation (14).

$$\begin{cases} \min \dfrac{1}{2}\|w\|^2 \\ s.t, y_i - wx_i - b \leq \varepsilon \\ wx_i - y_i + b \leq \varepsilon, i = 1, 2, 3..., n \end{cases} \qquad (14)$$

In equation (14), $\min \dfrac{1}{2}\|w\|^2$ can be used as a prerequisite for the optimization problem. When the current condition cannot be met, relaxation variables ξ_i and ξ_i^* are introduced to relax the range of the premise conditions. At this time, the objective function is to minimize the confidence range, as shown in equation (15).

$$\begin{cases} \min \dfrac{1}{2}\|w\| + C\sum_{i=1}^{N}\left(\xi_i + \xi_i^*\right) \\ s.t, y_i - wx_i - b \leq \varepsilon + \xi_i \\ wx_i - y_i + b \leq \varepsilon + \xi_i^*, i = 1, 2, 3...n \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases} \qquad (15)$$

In equation (15), C satisfies the penalty factor, which is the punishment for data that exceeds the ε range [16]. Through the above classification constraints, SVM has excellent ability in classification. The SVM is displayed in Figure 5.

In Figure 5, in neural network architecture, the input layer receives feature vectors from the original input of the data source, with each feature vector representing a sample. The hidden layer is composed of multiple neurons and connected to each input node in the input layer, forming a fully connected structure. The output layer is responsible for outputting classification results, and the number of neurons is usually equal to the number of categories in the classification problem. In the SVM structure, the input space refers to the space of the original input data, while the feature space refers to the space of the feature vectors obtained through hidden layer processing. SVM can map data from the input space to the feature space to complete classification tasks. The model structure for identifying damage in prestressed concrete components combining GABP and SVM is shown in Figure 6.

In Figure 6, the sensor of piezoelectric fluctuation measurement method identifies the fluctuation changes in prestressed concrete components. After collecting recognition data and conducting preliminary processing, it is input into a BP optimized by GA to identify the characteristics of concrete damage. Afterwards, the calculation results are fed into the SVM for further classification and correction, improving the accuracy of damage detection in prestressed concrete components. At the end, the visualization results of damage identification are output. A damage recognition method for prestressed concrete components is constructed on the basis of GABP-SVM. During the training, the original signal is first preprocessed. This includes removing the DC bias from the original signal to make the average value of the signal zero, and using a bandpass filter to remove high-frequency noise and low-frequency interference to preserve useful signal components. The most important

thing in the pre-processing process is to reduce dimensionality through SVD, retain the main singular values, and remove noise and redundant information. After that, a BP neural network is created and the dimensions of the input layer, hidden layer, and output layer are set. Then, GA is used to optimize the weights and thresholds of BP neural network to improve the performance of the model. Finally, SVM is used to classify the output of the optimized BP neural network to further improve the classification accuracy and complete

the model training. Compared with the traditional SVM, this study uses SVD to denoise the data. Moreover, the powerful nonlinear modeling capability of the improved BP algorithm is combined with the linear classification performance of SVM. By fully utilizing the advantages of these two algorithms, the classification problem of complex data collected by sensors can be better solved. This method enables GABP-SVM to better perform damage detection tasks.

(a) XGBoost recognition results

(b) AlexNet recognition results

- ○ Roller failure
- ○ Outer ring failure
- ○ Inner ring failure

(c) VGG16 recognition results

(d) GABP-SVM recognition results
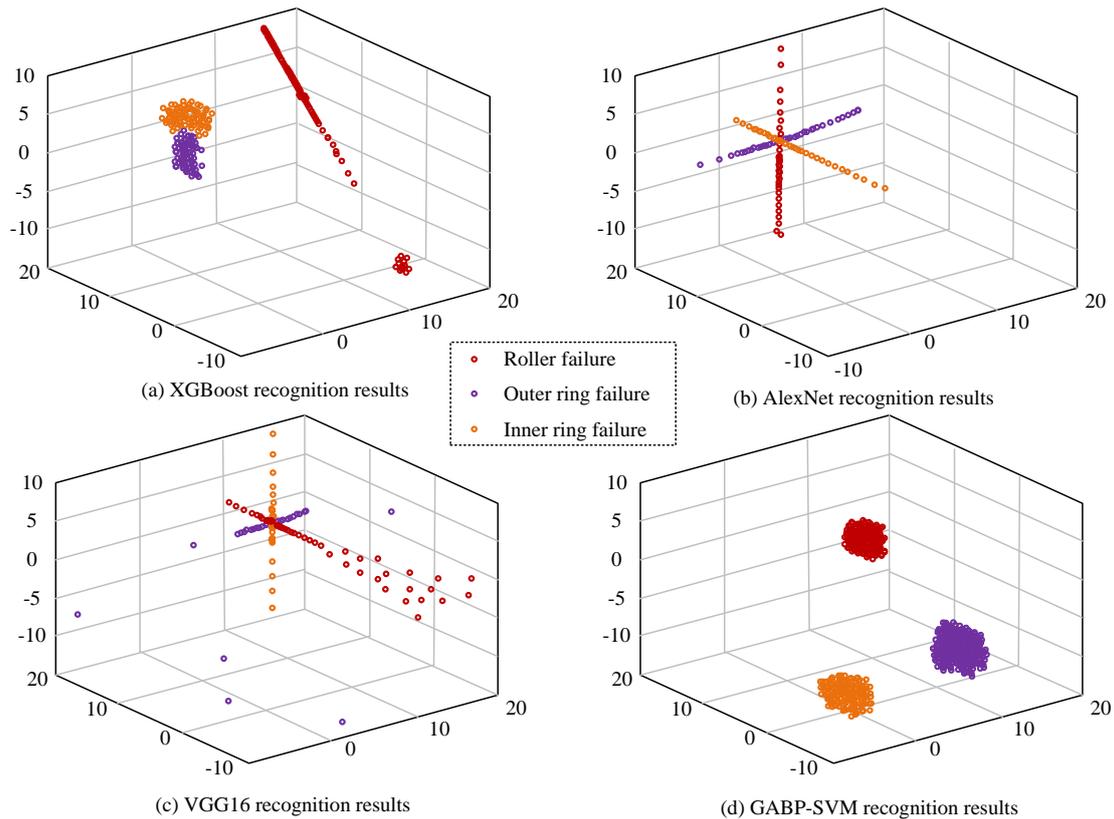
Figure 10: Comparison test results between GABP-SVM algorithm and other algorithms

Different combinations of mutation and crossover rates
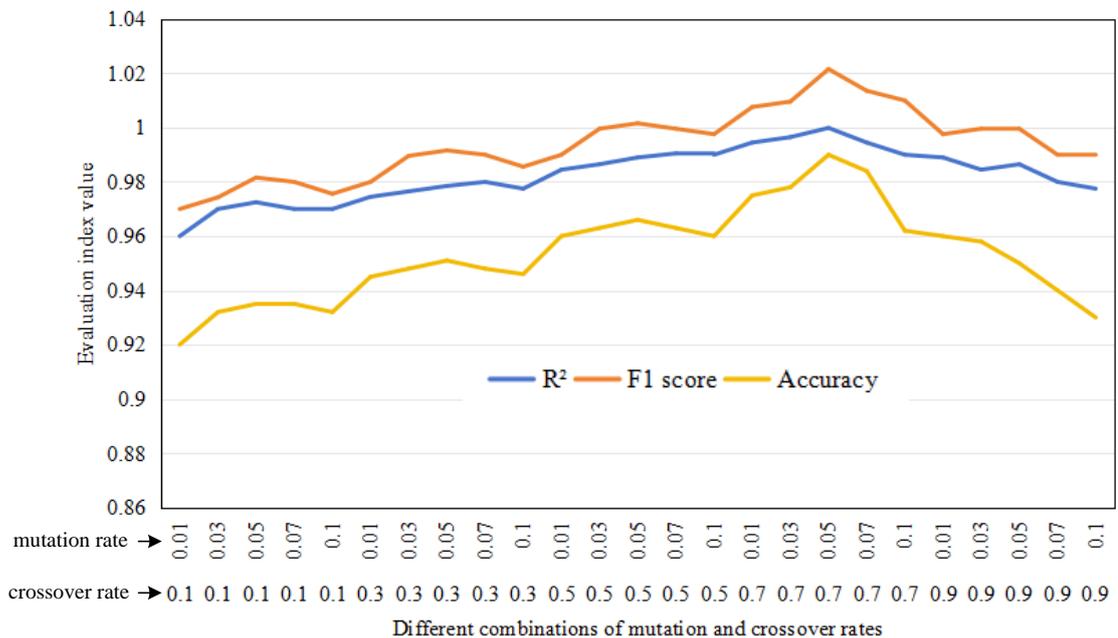
Figure 11: Comparison test results between GABP-SVM algorithm and other algorithms

(a) Damage location identification

(b) Damage degree identification
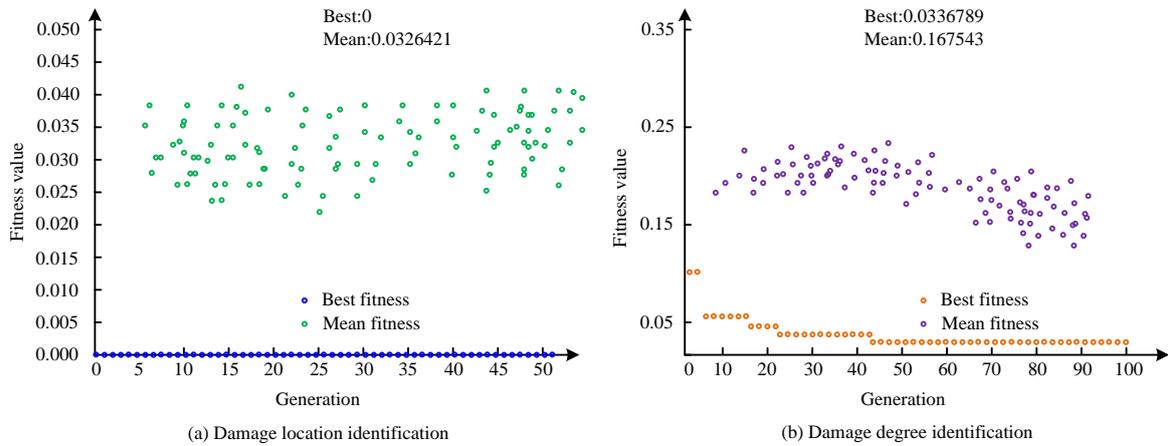
Figure 12: Experimental results of training GABP

(a) PSO-RF

(b) SVM
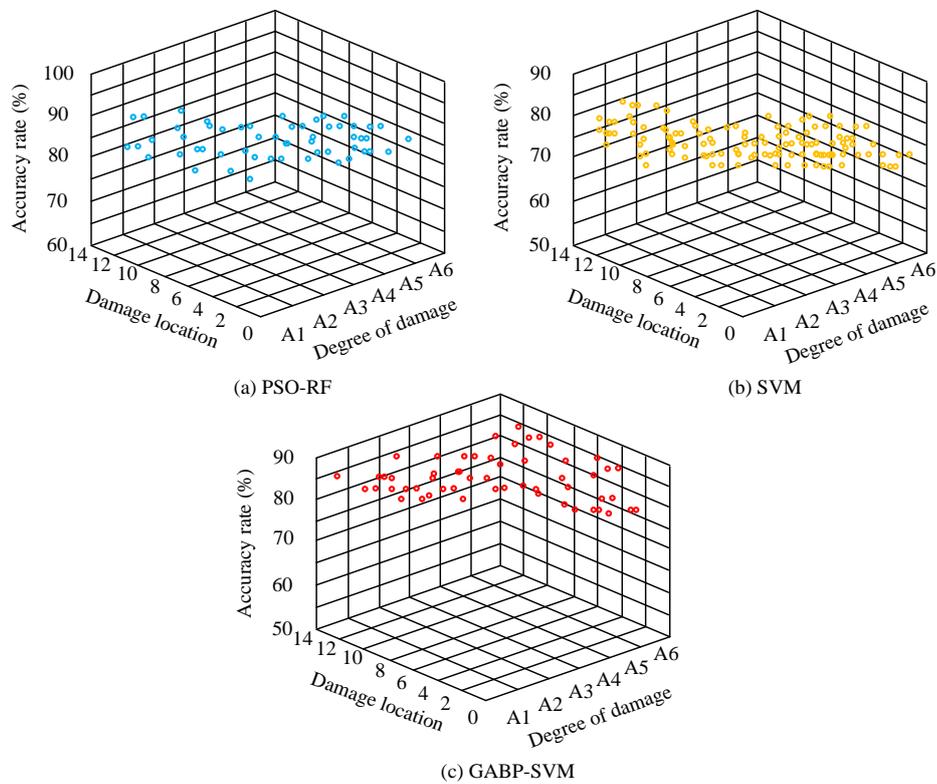
(c) GABP-SVM

Figure 13: Comparison of 4 algorithms for fault recognition in 3 parts

3 Results

To verify the damage identification model for prestressed concrete components based on GABP-SVM, relevant experiments are conducted. The experiment conducts performance tests on the optimization algorithm GA, feature classification algorithm GABP, and SVM used in the research, verifying the computational efficiency and classification accuracy. Afterwards, comparative tests are conducted with other concrete component damage identification models to verify the generalization ability and the damage identification performance.

3.1 Performance analysis of damage feature classification algorithm based on GABP-SVM

The study uses the piezoelectric wave measurement method to extract rich wave data information from prestressed concrete components. GA, BP, and SVM algorithms are used to identify damage characteristics. To verify the effect of BP optimized by GA, a comparative test is conducted between BP and GABP. The computer configuration used in the experiment is Intel ® CoreTM i9-9980XE, with an 8-core 2.1GHz CPU, 16GB of memory, and 1TB of hard drive. The

experiment selected 2,000 samples from the classic MNIST dataset for algorithm training and testing, with 80% of the dataset used for training and 20% for testing, as displayed in Figure 7.

In Figure 7 (a), the experimental results of the GABP algorithm on the testing set showed that the fitting degree between the GABP feature classification results and the true values was high. The R2 of the GABP was 0.9235, and the MSE was 0.00724. In Figure 7 (b), the BP algorithm on the testing set show that the fitting degree between the BP feature classification results and the true values was lower than that of GABP. The R2 of the BP was 0.8682, and the MSE was 0.01416. The results showed that the R2 value of GABP was 0.0553 higher than that of BP, and the MSE value of GABP was 0.00692 lower. The GABP has better feature classification ability than the basic BP algorithm. To verify the computational efficiency of the improved method, a comparative test is conducted between the GABP algorithm and three heuristic algorithms, namely Grey Wolf Optimization (GWO), Bat Algorithm (BA), and Particle Swarm Optimization (PSO), to optimize the BP algorithm, as displayed in Figure 8.

In Figure 8 (a), in the unimodal function, the average fitness of GABP dropped to the lowest at 417 iterations, and the average fitness of GWO-BP dropped to the lowest at 440 iterations. However, BA-BP algorithm and PSO-BP algorithm had slower iteration speeds than BA-BP and GWO-BP algorithms, and had not yet converged after 500 iterations, with average fitness values still at a high level. In Figure 8 (b), in multimodal functions, the average fitness value of the GABP algorithm dropped to the lowest after 48 iterations. The global search speed of the GWO-BP in the early stage was similar to that of the BA-BP algorithm, but the speed was slower in local fine search. The average fitness value of the GWO-BP algorithm dropped to the lowest after 174 iterations. The PSO-BP and BA-BP algorithms converged faster in multi-modal functions than in unimodal functions, but convergence was still incomplete after 500 iterations. The genetic operation of GA provides various ways to generate new solutions, increase the diversity of understanding, and help GA better explore the solution space. Therefore, GA has strong global search capabilities. Moreover, GA has strong parallel computing ability, which can significantly improve the convergence efficiency, so that the convergence efficiency of GA-BP is higher than that of PSO-BP and BA-BP. To verify the classification accuracy of SVM, a comparative test is conducted using Random Forest (RF), Decision Tree (DT), and SVM. 1,600 samples from the CIFAR-10 dataset were selected for algorithm training and testing, with 80% of the dataset used for training and 20% for testing. The results are shown in Figure 9.
In Figure 9 (a), the predicted value of SVM was close to the standard value, with a prediction accuracy of 0.94. In Figure 9 (b), the predicted value of the RF was far from the standard value, with a prediction accuracy of 0.74. In Figure 9 (c), compared with the standard value, the predicted value of DT was very discrete, and the distance between the predicted value and the standard value was

farther than SVM and RF, with a prediction accuracy of 0.58.

3.2   Performance Analysis of GABP-SVM Algorithm

To verify the performance of the GABP-SVM algorithm, experimental tests are conducted. The study conducted comparative experiments with Extreme Gradient Boosting (XGBoost) in ensemble learning, Visual Geometry Group Network (VGG), and AlexNet Neural Network (AlexNet). The experiment selected 20,000 data samples from the Kesi West Reserve bearing dataset, with 80% of the dataset used for model training and 20% for model testing. Four algorithms were used in the test to identify and classify faults in the inner, outer, and rolling elements of bearings. The additive Gaussian white noise in $-8 \le SNR(db) \le 3$ was introduced as interference data. The experimental results are shown in Figure 10.

In Figure 10 (a), XGBoost algorithm identified the fault characteristics of the rolling elements, but the fault characteristics of the inner and outer rings were not accurately classified. In Figure 10 (b), AlexNet algorithm could recognize the inner, outer, and rolling element features, but did not complete the classification task for the three types of features. In Figure 10 (c), VGG16 algorithm did not complete the recognition and classification of fault types, and the recognition results had many discrete points. The accuracy and robustness of the model were both at a low level. In Figure 10 (d), GABP-SVM algorithm accurately identified the fault characteristics of the inner and outer rings and rolling elements of rolling bearings and completed the classification. The results show that in noisy environments, the feature recognition ability of GABP-SVM algorithm is superior to XGBoost, AlexNet, and VGG16 algorithms. The setting of crossover and mutation rates in GA will affect the optimization effect of parameters, thereby affecting the performance of the model. To analyze the impact of GA parameters on model performance, a sensitivity analysis is conducted. The study conducted tests different combinations of mutation rate and crossover rate, and the test results are shown in Figure 11.

In Figure 11, as the crossover rate increased, the performance of the model gradually improved. When the crossover rate was 0.7, the R2, F1 score, and accuracy of the model were all the highest. When the crossover rate was 0.9, the indicator values decreased. In different crossover rate values, the indicator value first increased and then decreased with the increase of mutation rate. When the crossover rate was 0.7 and the mutation rate was 0.05, the R2, F1 score, and accuracy were at their highest values, which was the optimal parameter setting for GA in GABP-SVM.

3.3   Performance analysis of damage identification model for prestressed concrete components based on GABP-SVM

The above experiment shows that the feature classification algorithm used in the study has high accuracy. To verify the performance of combining the two for damage feature recognition, a comparative experiment is conducted between the GABP-SVM model and concrete damage recognition models based on SVM and GABP, as well as a model that integrates PSO and RF. The experiment first trains GABP to obtain the optimal weights and thresholds of BP, as presented in Figure 12.

In Figure 12 (a), after 54 iterations, the calculated optimal fitness value was 0.0326421, and the optimal number of hidden layers was 9. In Figure 12 (b), after 100 iterations, the optimal fitness value calculated was 0.0336789, and the optimal hidden layer was 21. The BP is set as the optimal threshold and weights. A comparative experiment is conducted between the optimized GABP-SVM and other models mentioned above, as presented in Table 2.

In Table 2, the symbol "*" indicates a P<0.05 when compared to other algorithms, indicating that the difference in results is statistically significant. Among them, R2 is mainly used for regression analysis to determine the classification ability of the model. The range of R2 values is [0, 1], and the larger the R2, the higher the classification accuracy of the model. The Mean Squared Error (MSE), R2, and F1 scores of the GABP-SVM model in the damage location recognition results were 7.962×10^{-4}, 0.9756, and 0.9836, respectively, which were superior to the SVM model, GABP model, and PSO-RF model. In the results of identifying the degree of damage, the MSE, R2, and F1 scores of the GABP-SVM model were 6.548×10^{-2}, 0.9531, and 0.9925, respectively, which were also superior to other models. To further verify the robustness, comparative tests are conducted on the models that performed better in an environment with added noise. The test results are shown in Figure 13.

In Figure 13 (a), the PSO-RF model had an accuracy of 82.6% in identifying damage to prestressed concrete components. In Figure 13 (b), the accuracy of SVM model in identifying damage to prestressed concrete components was 69.8%. In Figure 13 (c), the accuracy of GABP-SVM model for damage identification of prestressed concrete components was 93.7%. The results indicate that in noisy environments, the damage recognition accuracy of the GABP-SVM model is significantly higher than that of the SVM and PSO-RF models. To further analyze the robustness of the research model under different damage scales, the experimental data is organized into three datasets: minor damage, moderate damage, and serious damage to test the model. The experimental results are shown in Table 3. In Table 3, "*" represents a P<0.05 when compared to other algorithms, indicating that the difference in results is statistically significant. The recognition ability of

GABP-SVM and PSO-RF models for serious, moderate, and minor damages showed a decreasing trend, while the decrease in GABP-SVM was smaller than that of PSO-RF. GABP-SVM maintained a high overall recognition level, with MSE, R2, and F1 scores of 0.0907, 0.5054, and 0.4062 for PSO-RF for minor damages, respectively. The results indicate that the GABP-SVM model exhibits higher robustness than the comparison model under different damage scales.

4   Discussion

Comparative experiments were conducted on the basic algorithms that make up the research model, and a damage identification model for prestressed concrete components based on GABP-SVM was verified. The R2 of the GABP feature recognition was 0.9235, and the MSE value was 0.00724, which was better than the BP algorithm. Jin et al. reached similar conclusions using GABP for rolling bearing fault diagnosis [17]. In the calculation of unimodal and multimodal functions, the GABP algorithm solved for the optimal solution that was closer to the true result after 417 and 48 iterations, respectively. This result was similar to the one obtained by Khatri et al. using GA to improve the load-bearing performance of fluid dynamic sliding bearings [18]. However, the research algorithm exhibits higher feature recognition ability and convergence efficiency compared to the latest advanced algorithms. The reason is that the GA is used to automatically optimize the parameters of the BP algorithm. In addition, the GA is improved to avoid getting stuck in local optima. The classification prediction accuracy of SVM algorithm was 0.94, which was better than RF algorithm and DT algorithm. This result was similar to the conclusion of the SVM-based radial deformation error evaluation method for turbine blades proposed by Chen et al [19]. In the model testing, the GABP-SVM model performed better than the SVM model, GABP model, and PSO-RF model in identifying the location and degree of damage. In the robustness test of the model with added noise, the accuracy of the GABP-SVM model in identifying damage to prestressed concrete components was 93.7%, which was 11.1% and 23.9% higher than the PSO-RF and SVM. Zhao et al. proposed similar conclusions in the research of concrete mesoscopic damage characteristics detection based on improved R-CNN [20]. However, the GABP-SVM model is more robust than the newly proposed model based on the improved R-CNN model. Also, due to the deep optimization of GA in this study, it avoids falling into local optimal solutions in complex environments, thus ensuring the stability of the model. The results indicate that the designed model not only has excellent accuracy in identifying damage to concrete components, but also has extremely high robustness. From this, the study uses RF to preprocess the data and GA to optimize the parameters of BP. The improved BP output results are input into SVM for further feature classification, greatly improving the recognition ability of concrete component damage and the generalization ability in complex concrete structures. Finally, the study analyzes

the computational complexity and scalability of the GABP-SVM model. The analysis results show that as the data size increases, the computational complexity does not significantly increase. In high-dimensional data, the GABP-SVM model can still maintain high feature recognition ability. The results indicate that the dimensionality reduction techniques introduced in the study and the parallel computing capability of the GA enable the GABP-SVM model to have higher computational efficiency and scalability.

5 Conclusion

Concrete is widely used in urban construction and industrial production. Concrete components are prone to damage and structural instability under conditions such as earthquakes, long-term high loads, and environmental corrosion. A smart concrete damage identification model was constructed by combining machine learning algorithms with the intelligent sensing effect of piezoelectric materials, aiming to accurately evaluate the health status of concrete. The study used piezoelectric wave measurement method to collect small wave signals from concrete. The GA was used to optimize BP to identify the characteristics of concrete damage from the signals. In addition, SVM was introduced to further classify and modify the recognition results of GABP, and a damage recognition model for prestressed concrete components based on GSBP-SVM was constructed. The designed model could accurately identify the damage location and degree of concrete, with high robustness. Concrete has extremely wide applications. Currently, there are a large number of buildings with complex structures and large volumes. In addition to prestressed concrete components, various forms of concrete components are different. Future research can focus on different forms of concrete components to broaden the applicability of the research model.

Funding

References

[1] Gangadhar Bandewad, Kunal P. Datta, Bharti W. Gawali, and Sunil N. Pawar. Review on discrimination of hazardous gases by smart sensing technology. Artificial intelligence and applications, 1(2):86-97, 2023. https://doi.org/10.47852/bonviewAIA3202434

[2] Qianqian Wang, Aimin An, Minan Tang, and Jiawei Lu. Distributed nonlinear model predictive control for cobalt removal process in zinc hydrometallurgy considering error compensation modelling. The Canadian Journal of Chemical Engineering, 102(1):307-323, 2024. https://doi.org/10.1002/cjce.25036

[3] Long Bai, Xin Cheng, Qizhong Yang, and Jianfeng Xu. Predictive model of surface roughness in milling of 7075Al based on chatter stability analysis and back propagation neural network. The International Journal of Advanced Manufacturing Technology, 126(3):1347-1361, 2023. https://doi.org/10.1007/s00170-023-11133-6

[4] Fazhan Zeng, Ren Wan, Yongjun Xiao, Fan Song, Changjun Peng, and Honglai Liu. Predicting the self-diffusion coefficient of liquids based on backpropagation artificial neural network: a quantitative structure-property relationship study. Industrial & Engineering Chemistry Research, 61(48):17697-17706, 2022. https://doi.org/10.1021/acs.iecr.2c03342

[5] Jiang L, Duan J J, Zheng R P, Shen H N, Li H, and Xu J. Optimization and simulation of garment production line balance based on improved GA. International Journal of Simulation Modelling, 22(2):303-314, 2023. https://doi.org/10.2507/ijsimm22-2-co6

[6] Xinzhe Yin, Jinghua Li, and Shoujun Huang. The improved genetic and BP hybrid algorithm and neural network economic early warning system. Neural computing & applications, 34(5):3365-3374, 2022. https://doi.org/10.1007/s00521-021-05712-5

[7] E. Gangadevi, R. Shoba Rani, Rajesh Kumar Dhanaraj, and Anand Nayyar. Spot-out fruit fly algorithm with simulated annealing optimized SVM for detecting tomato plant diseases. Neural computing & applications, 36(8):4349-4375, 2024. https://doi.org/10.1007/s00521-023-09295-1

[8] Dongmei Xu, Xiangqi Wang, Wen Wang, K. Chau, and Hongfei Zang. Improved monthly runoff time series prediction using the SOA-SVM model based on ICEEMDAN-WD decomposition. Journal of hydroinformatics, 25(3/4):943-970, 2023. https://doi.org/10.2166/hydro.2023.172

[9] Hongchang Zhang, Yang Zou, Enrique del Rey Castillo, and Xiaofei Yang. Detection of RC spalling damage and quantification of its key properties from 3D point cloud. KSCE journal of civil engineering, 26(5):2023-2035, 2022. https://doi.org/10.1007/s12205-022-0890-y

[10] Xing Fan, Mengdi Sun, Zhong Li, Zhenhua Chen, Xiaoyong Zhou, Quanxuan Lu, and Zhicheng Zhang. Tuning piezoelectric properties of P(VDF-TFE) films for sensor application. Reactive and functional polymers, 180(1):50-57, 2022. https://doi.org/10.1016/j.reactfunctpolym.2022.105391

[11] Nasim Kamely. Interaction of light with different electroactive materials: a review. Journal of electronic materials, 51(3):953-965, 2022. https://doi.org/10.1007/s11664-021-09362-0

[12] He J, and Yang J. Network security situational level prediction based on a double-feedback elman model. Informatica: An International Journal of

Computing and Informatics, 46(1):87-93, 2022. https://doi.org/10.31449/inf.v46i1.3775

[13] Kumar Y, Dahiya N, Malik S, and Savita K. A new variant of teaching learning-based optimization algorithm for global optimization problems. Informatica, 43(1):65-75, 2019. https://doi.org/10.31449/inf.v43i1.1636

[14] Panda D, Panda D, Dash S R, and Shantipriya P. Extreme learning machines with feature selection using GA for effective prediction of fetal heart disease: a novel approach. Informatica: An International Journal of Computing and Informatics, 45(3):381-392, 2021. https://doi.org/10.31449/inf.v45i3.3223

[15] Rui Wang, Anxiang Song, Xiang Chen, Yuanchen Guo, Xue Wang, Yan Sun, and Miao Tian. Experimental study on the properties of phosphate-based materials for rapid repair of concrete cracks. KSCE Journal of Civil Engineering, 26(5):2342-2353, 2022. https://doi.org/10.1007/s12205-022-1192-0

[16] Kulkarni S A, Gurpur V, Koval K A. Impact of Gaussian Noise for Optimized Support Vector Machine Algorithm Applied to Medicare Payment on Raspberry Pi. Informatica: An International Journal of Computing and Informatics, 45(4):643-652, 2021. https://doi.org/10.31449/inf.v45i4.3747

[17] Tongtong Jin, Qiang Cheng, Hu Chen, Siyuan Wang, Jinyan Guo, and Chuanhai Chen. Fault diagnosis of rotating machines based on EEMD-MPE and GA-BP. The International Journal of Advanced Manufacturing Technology, 124(11):3911-3922, 2023. https://doi.org/10.1007/s00170-021-08159-z

[18] Chandra B. Khatri, Saurabh K. Yadav, Gananath D. Thakre, and Arvind K. Rajput. Design optimization of vein-bionic textured hydrodynamic journal bearing using genetic algorithm. Acta mechanica, 235(1):167-190, 2024. https://doi.org/10.1007/s00707-023-03734-9

[19] Junyu Chen, Yun Feng, D. Teng, and Cheng Lu. Support vector machines-based pre-calculation error for structural reliability analysis. Engineering with computers, 40(1):477-491, 2024. https://doi.org/10.1007/s00366-023-01803-0

[20] Liang Zhao, Shenglun Gao, Junying Chen, and Jiajia Li. Feature detection of concrete mesoscopic damage based on feature sharing double-head Cascade R-CNN. Control and Decision, 37(7):1-7, 2022. https://doi.org/10.13195/j.kzyjc.2021.0124