Damage identification of composite based on moving principal component analysis and support vector machine

Liangliang Zhou [1], Yang Shen [1], Hongxiang Jiang [1],
Xu Zeng [2], Yinan Shan [1, 3, *]

[1] Zhan Tianyou Honors College, Dalian Jiaotong University, 116028, China

[2] State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology

[3] CRRC QINGDAO SIFANG CO., LTD., Qingdao 266031, Shandong, China;

* Corresponding author's e-mail: shanyn@djtu.edu.cn

Abstract. The vibration response of composite under dynamic excitation is used for structural condition monitoring and damage identification, which has important engineering application value. In this paper, a Particle Swarm Optimization based Support Vector Machine (PSO-SVM) damage identification method is used to classify and predict the location and degree of volume loss damage of Carbon Fiber Reinforced Polymer (CFRP). Through numerical simulation and experimental verification, the feature parameters were extracted by principal component analysis, and then the number of input feature parameters was optimized by PSO-SVM, and the damage was classified and predicted. In this paper, the classification performance of several common Machine Learning (ML) models is compared. The results show that the average accuracy of the PSO-SVM model is 97.22% and 81.94%, both higher than that of Back Propagation Neural Network (BP) and Particle Swarm Optimization based Back Propagation Neural Network (PSO-BP).

Keywords: CFRP; Damage identification; Machine learning; PSO-SVM; MPCA.

1. Introduction

CFRP is widely used because of its superior performance. However, the carbon fiber composite structure is prone to damage in the process of preparation and service, and some damage is not easy to be found in the early stage, with high concealment, which is a great harm to the service safety of the structure. Therefore, it is very important to find a method that can carry out structural health monitoring to ensure the safety and reliability of composite structures.

In recent years, domestic and foreign scholars have used some computational methods to predict the damage degree and severity of structures based on the dynamic response of structures [1-4]. In addition, Diao used Hilbert spectrum energy to construct structural damage characteristics, and used support vector machine to detect and experiment on the location and degree of damage [5]. Khatir combined PSO and Jaya algorithm with extended finite element (XFEM) and extended geometric analysis (XIGA) to predict crack location [6]. Yu proposed a support vector machine multi-classifier optimized based on genetic algorithm to realize effective and accurate recognition of the working type of wooden poles [7]. Cuong-Le proposed a damage identification method combining PSO and SVM, The result shows that compared with other machine learning models, the proposed PSO-SVM has higher prediction accuracy in terms of damage location and damage degree [8].

However, the ML model still has some problems, such as being easily disturbed by noise, relying on artificial features, and failing to meet the accuracy requirements of damage identification. In view of the problem that existing features cannot meet the accuracy of damage recognition, principal component analysis method can realize data compression and extract data features to meet the problem of insufficient features and existing features cannot meet the accuracy of damage recognition. In this paper, MPCA optimization features are used as the input of PSO-SVM model to predict the damage location and damage grade of composite plates. The results show that the average accuracy of PSO-SVM model is 97.22% and 81.94% respectively, both higher than that of BP model and PSO-BP model.

2. Methodology

2.1 Feature extraction method

PCA can compress the multi-dimensional features of the original data into a few comprehensive features, which is a mathematical dimensionality reduction method. After dimensionality reduction, each feature vector is linearly independent, and the original data information can be preserved to the maximum. Based on PCA, MPCA introduces the concept of time window to obtain the eigenvector time series after PCA.

When the plate structure is damaged, the strain response changes and the eigenvector changes accordingly. The relative variation of the first principal component eigenvector of composite plate structure, The first principal component eigenvector modulus length and the directional angles of the eigenvector variation (DAEV) can be used as new eigenvectors in the monitoring process. The specific expression is as follows:

$$\Delta\psi_1(k)=\psi_1(k)-\overline{\psi}_1^{h} \tag{1}$$

$$\left|\Delta\psi_1(k)\right|=\sqrt{\sum_{j=1}^{M}\Delta\psi_{1j}^2(k)} \tag{2}$$

$$\alpha_{1j}(k)=\arccos\left(\frac{\Delta\psi_{1j}(k)}{\left|\Delta\psi_1(k)\right|}\right) \tag{3}$$

where $\overline{\psi}_1^{h}$ is the mean of the first principal eigenvector of the health state, $\psi_1(k)$ is the first principal eigenvector under the k analysis step, $\Delta\psi_{1j}(k)$ is the relative change between the eigenvector component $\psi_{1j}(k)$ and the mean value of the first eigenvector component in the structural health state at the k analysis step.

In addition, the relative entropy index can be obtained by the principal component analysis of the strain data after the energy decomposition of the wavelet packet. That is, the principal component score matrix of the energy strain data obtained from different measuring points and the data from other measuring points can be obtained by the principal component analysis, and the relative entropy value (RE) of the matrix can be extracted as the local damage characteristics of the measuring point.

$$WES_{i,j}=\sqrt{\frac{\sum_{k=1}^{N}\varepsilon_{i,j,k}^2}{N}}=N^{-1/2}\left\|\varepsilon_{i,j}\right\|_2$$

(4)

where $WES_{i,j}$ represents the wavelet packet energy strain in the j band of the measuring point i, k represents the response data point, and N represents the sampling number in this period.

2.2 Damage identification framework based on PSO-SVM

PSO is a robust and efficient algorithm that constantly updates the position and velocity of each particle to find an optimal solution. The core of SVM model is to determine the weight vector and scalar threshold of the mathematical structure of the model, so that the error of the model expression in the data set is minimized, and the maximum margin is taken as the most effective classification solution. Basically, the ideas are generalized in the following equation.

$$\begin{cases} f(x) = \omega \cdot x - b \\ \underset{\omega,b,\varepsilon}{Minimize} \frac{1}{2}\|\omega\|^2 \\ \left|\left(\langle \omega x_i \rangle + b\right) - y_i\right| \leq \varepsilon \end{cases}$$

(5)

where x is the input vector, ω is weight vector, b is scalar threshold, $\varepsilon \geq 0$ is error-insensitive zone and $\|\omega\|^2 = \langle \omega \cdot \omega \rangle$ is the norm of the weight vector.

Firstly, the original strain signal is decomposed into different frequency components by wavelet packet energy transformation, and the energy information at each frequency component is obtained to form the wavelet packet energy vector. Then, the wavelet packet energy signal and the original signal are reduced by principal component analysis and the main features are extracted. The damage mode of the composite plate is identified based on the relative entropy and the feature vector modulus length. The selected features are taken as the input of the training model, SVM classifier is established, and the recognition results are output. In order to improve the recognition accuracy of the classifier, PSO algorithm was applied to optimize the model parameters, and the model was tested by randomly extracting data according to the damage location and damage grade.

Table 1 shows the model impact factors of different machine learning models for analysis.

Table 1. General model factors applied.

Model name	Model factors
BP	Hidden layer nodes=12, training times=1000, learning rate=0.01
PSO-BP	C1=2, C2=2, population =12, hidden layer nodes=12, training times=1000, learning rate=0.01
PSO-SVM	C1=1.5, C2=2.5, population=12, maximum Iteration=200, Radial basic kernel function

3. Experimental study and result discussion

3.1 Numerical example and results

Taking composite material board as an example, finite element simulation platform was used for simulation, and its parameters were set as shown in Table 2 and Table 3.

Table 2. Finite element simulation setup.

Category	Specific parameter
Model size	400mm×300mm×2mm
Lay-up mode	[0°/90°/0°/90°/0°/90°/0°/90°/0°/90°/0°]
Boundary condition	One narrow side is fixed, the other side is free
Cell type	Shell281
Unit partition	Divide 40 units in the x direction and 30 units in the ydirection
Damage element coordinate	Unit ① (17,6), Unit ② (28,6), Unit ③ (14,14), Unit ④ (22,15), Unit ⑤ (28,17), Unit ⑥ (17,24)
Loading load	Force hammer load, as shown in Figure 2

The damage simulation of composite materials is based on the stress description, and both the additional mass and the through hole can cause local stress concentration, and both can lead to local stiffness reduction.

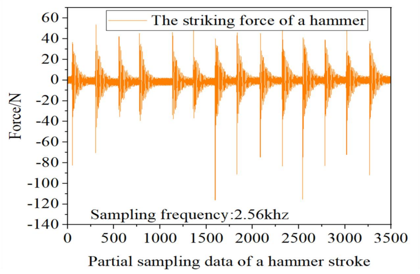


Fig 1. Sampling data of partial percussion force of a force hammer

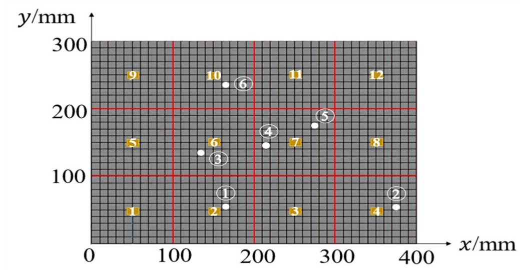
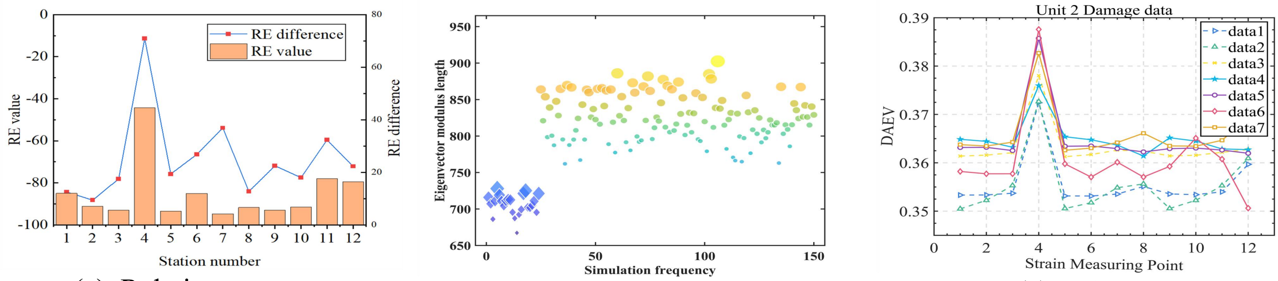


Fig 2. Simulation model and numbering Settings

The simulated strain data were used to construct datasets M1 and M2 with a sampling frequency of 2.56khz, 12 sampling points, and each sampling time of 5s. M1 data set is in lossless working condition, M2 data set is in lossy working condition.

Fig. 3 respectively shows the changes of three characteristic indexes of each measuring point under the damage state of unit ④. It can be seen that the damage unit ④ can cause the mutation of the three indicators at the measuring point of the area where the damage is located, so as to identify whether there is damage in the area. Among them, (a) The figure shows that both the RE value and the RE difference have mutations at test point 4, At the same time (b) Figure shows that the modulus length of the feature vector difference vector increases after the damage occurs, and there is an obvious threshold, (c) It can be seen that DAEV index can effectively distinguish damaged units.



(a) Relative entropy (b) Eigenvector modulus length (c) DAEV

Fig 3. Simulation data verification results of three indexes

3.2 Test and results

During the test, the composite plate is randomly excited by the exciter, and the strain data of 12 measuring points are finally collected by the DH5922N signal acquisition system through the adapter, the frequency sweep signal generator and signal amplifier. The schematic diagram of the test device is shown in Fig 4. The input database was composed of 240 sets of data containing random damage locations and damage grades. To avoid overfitting of the model, the database was randomly divided into 168 groups for training and 72 groups for testing. The damage rating is divided into seven levels, as shown in Table 3 below. The damage locations were evenly distributed among 12 units. Each set of data contains a total of 37 eigenvector values of the three indicators, in which RE contains 12 measurement point values, DAEV contains 12 measurement point values, and the difference vector modulus length contains 13 values.

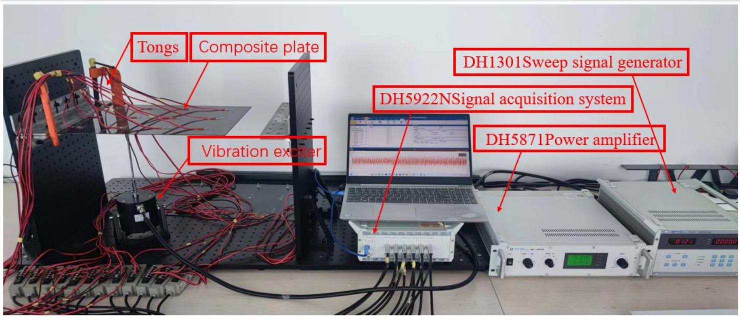
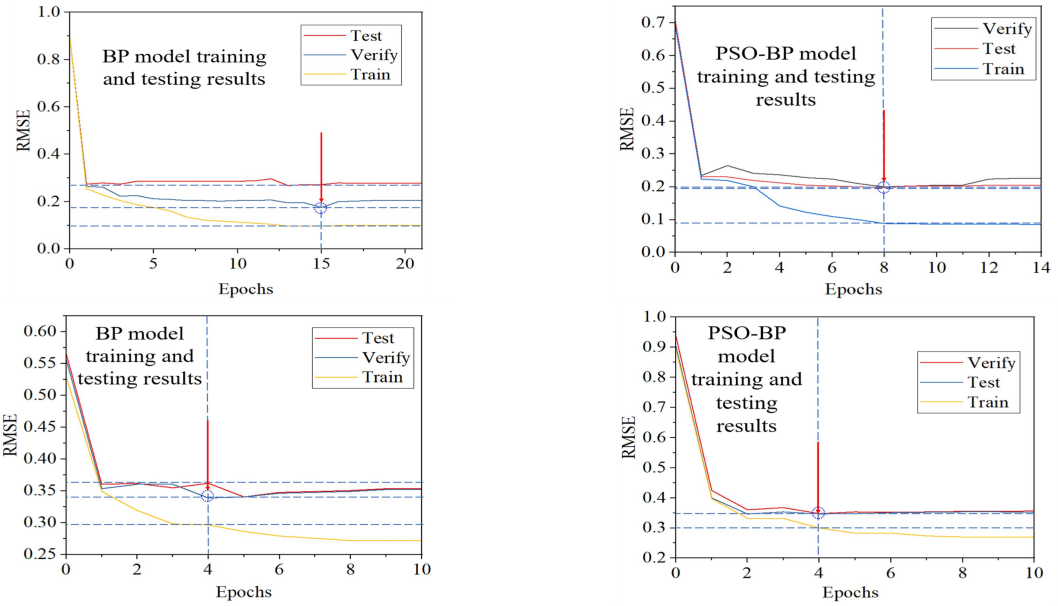


Fig 4. Test apparatus

Table 3. Mass block quality grade

Mass block grade	I	II	III	IV	V	VI	VII
Mass block mass/g	3.56	4.62	5.58	7.22	10.28	12.48	14.26

Different ML models such as BP, PSO-BP and PSO-SVM are trained for classification and prediction. The results of training and testing using BP and PSO-BP neural network are shown in Fig 5. The optimal cycle of BP neural network is 15, and the optimal cycle of PSO-BP is 8. PSO-BP has a better performance than BP.



(a) Results of BP model based on Composite plate
(b) Results of PSO-BP model based on Composite plate

Fig 5. Training and testing BP Neural Network and PSO-BP for the Composite plate

Fig 6 plots several different input features of PSO-SVM (12, 24, 25, 36, 37) and their performance. The optimal number of features is 24, and the performance of input parameters (37 features) is worse than that of 24 input features. The analysis shows that particle swarm optimization can effectively eliminate redundant parameters and noise, and improve the performance of the model through powerful search function.

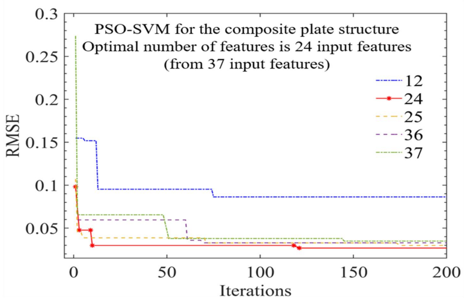


Fig 6. Training with the PSO-SVM for the Composite plate

Table 4 and Table 5 show a comparison between the classification results of these machine learning models. As can be seen, PSO-SVM shows higher accuracy in these machine learning models. The accuracy of BP and PSO-BP for damage location classification in the test set were 73.6111% and 79.1667%, and RMSE were 0.2846 and 0.19493, respectively. It can be seen that the performance of the two neural network models is not ideal. Meanwhile, the accuracy of the above two methods for judging the damage degree of the test set is 61.112% and 69.444%, and the RMSE is 0.36221 and 0.34836, respectively. In contrast, the accuracy of the PSO-SVM model for the classification prediction of damage location and damage grade of the test set is 97.222% and 81.944%, respectively, which is higher than other ML models.

Table 4. Comparison of training and testing results of different ML models for damage locations.

Model	BP		PSO-BP		PSO-SVM	
	RMSE	Acc	RMSE	Acc	RMSE	Acc
Training set result	0.09382	0.91071	0.08831	0.95238	0.05357	0.99404
Test set result	0.28460	0.73611	0.19493	0.79166	0.08285	0.97222

Table 5. Comparison of training and testing results of ML models with different damage levels.

Model	BP		PSO-BP		PSO-SVM	
	RMSE	Acc	RMSE	Acc	RMSE	Acc
Test set result	0.36221	0.61111	0.34836	0.69444	0.18381	0.81944

4. Summary

Based on the results and discussions presented above, the conclusions are obtained as below:

(1) The optimal number of features of the composite structure is 24/37, indicating that the input database contains a large number of redundant or noisy parameters. The proposed PSO-SVM can effectively determine the optimal input features and eliminate redundant and noisy data.

(2) Compared with BP and PSO-BP models, PSO-SVM model showed excellent performance, mainly in the classification and prediction of damage location and damage degree (the accuracy of classification and prediction of damage location and damage grade were 97.22% and 81.94%, respectively), while other ML models were not ideal in the prediction of damage degree.

Acknowledgments

This work was financially supported by the Fundamental Research Funds for the Provincial Universities of Liaoning (Project No. LJ212410150031).

References

[1] Silva, M., et al., Deep principal component analysis: An enhanced approach for structural damage identification. Structural health monitoring, 2019. 18(5-6): p. 1444-1463.

[2] Zhang, G., et al., Machine-learning-based damage identification methods with features derived from moving principal component analysis. Mechanics of advanced materials and structures, 2020. 27(21): p. 1789-1802.

[3] Zhu, Y., et al., A temperature-driven MPCA method for structural anomaly detection. Engineering structures, 2019. 190: p. 447-458.

[4] Bisheh, H.B. and G.G. Amiri, Structural damage detection based on variational mode decomposition and kernel PCA-based support vector machine. Engineering Structures, 2023. 278: p. 115565.

[5] Diao, Y., et al., Structural damage identification using modified Hilbert–Huang transform and support vector machine. Journal of Civil Structural Health Monitoring, 2021. 11(4): p. 1155-1174.

[6] Khatir, S. and M. Abdel Wahab, A computational approach for crack identification in plate structures using XFEM, XIGA, PSO and Jaya algorithm. Theoretical and applied fracture mechanics, 2019. 103: p. 102240.

[7] Yu, Y., et al., Wavelet packet energy–based damage identification of wood utility poles using support vector machine multi-classifier and evidence theory. Structural Health Monitoring, 2019. 18(1): p. 123-142.

[8] Cuong-Le, T., et al., An efficient approach for damage identification based on improved machine learning using PSO-SVM. Engineering with computers, 2022. 38(4): p. 3069-3084.