Article

Energy Ratio Variation-Based Structural Damage Detection Using Convolutional Neural Network

Chuan-Sheng Wu 1, Yang-Xia Peng 2, De-Bing Zhuo 3,*, Jian-Qiang Zhang 2, Wei Ren 2 and Zhen-Yang Feng 4

1 School of Economics and Management, Chongqing Jiaotong University, Chongqing 400074, China
2 School of Civil Engineering, Chongqing Jiaotong University, Chongqing 400074, China
3 School of Civil Engineering and Architecture, Jishou University, Zhangjiajie 427000, China
4 Housing and Urban and Rural Construction Commission, Banan District, Chongqing 401320, China
* Correspondence: zhuodebing@163.com

Abstract: In the field of structural health monitoring (SHM), with the mature development of artificial intelligence, deep learning-based structural damage identification techniques have attracted wide attention. In this paper, the convolutional neural network (CNN) is used to extract the damage feature of simple supported steel beams. Firstly, the transient dynamic analysis of the steel beam is carried out by finite element software, and the acceleration response signals under different damage scenarios are obtained. Then, the acceleration response signal is decomposed by wavelet packet decomposition (WPD) to extract the wavelet packet band energy ratio variation (ERV) index as the training sample of CNN. Subsequently, the vibration experiment of a simple supported steel beam was carried out, and the results were compared with the numerical simulation results. The characteristic indexes were obtained by making corresponding changes to the vibration signal, and then, the experimental data were input into the CNN to predict the effect of damage detection. The results show that the method can successfully detect the intact structure, single damage, and multiple damages with an accuracy of 95.14% under impact load, and the performance is better than that of support vector machine (SVM), with good robustness.

Keywords: simple supported steel beams; structural damage detection; convolutional neural network; wavelet packet decomposition; energy ratio variation

1. Introduction

Structural damage detection is an important research field of structural health monitoring (SHM) [1–3]. During the usage of a bridge structure, the internal damage to the structure will be caused by material corrosion and structural performance degradation. In order to prevent casualties and property losses caused by the sudden collapse of the structure, it is urgent to carry out the identification of structural damage. Traditionally, whether the structure is damaged is judged by changes in the time domain, frequency domain, and mode. It is usually necessary to fix the sensor to obtain the vibration signal of the structure. The obtained signal is transformed into the corresponding identification index through algorithm processing, so as to qualitatively judge the damage location of the structure and quantitatively judge the damage degree of the structure. Among them, the damage identification method of global dynamic characteristics based on natural frequency [4,5], modal shape [6], and modal strain energy [7] (MSE) is not easy to observe obvious damage through experiments in the process of determining the damage location and damage degree, and the natural frequency is a global variable. It is difficult to identify the local damage of the structure, especially to judge the damage of the symmetrical position under the load. Although the MSE can identify the damage location of the unit, the complete modal shape is needed in the process of solving, so there are also some problems in the extension of the vibration mode, and there will be deviations in the actual measurement process, which is difficult to apply to the practical project.
In order to overcome the above limits, a signal processing technology, such as wavelet analysis, is proposed, which is widely used in the field of structural damage detection [8–10]. Then, wavelet packet analysis, which can decompose low-frequency and high-frequency signals simultaneously, is introduced. It has higher decomposition precision and a better signal processing ability [11–13]. The extracted damage characteristics include the wavelet coefficient [14], wavelet or wavelet packet band energy [15,16], wavelet energy ratio [17], etc. Since the energy before and after structural damage will be partially suppressed, so the constructed energy ratio variation (ERV) can correspond to the damage of the structure, which can better reflect the damage information of the structure. In the process of extracting features, it will be disturbed by some noise, so it is necessary to further process the features. For the traditional method of the artificial neural network (ANN) [18,19], the most representative are the BP neural network [20] and radial basis function (RBF) [21,22] neural network, which can not only extract the corresponding eigenvalues but also reduce the noise interference.

Support vector machine (SVM) and ANN, as two highlights of damage identification methods, are used by many scholars to deal with the problems of structural damage and classification. However, due to their inherent defects, such as only dealing with small sample problems, the calculation process is time-consuming, which easily leads to overfitting of data and has a great impact on early damage detection [23]. Therefore, in order to overcome the shortcomings of traditional neural network algorithms and machine learning algorithms, this paper introduces a convolutional neural network (CNN) that can process a large amount of data and does not require manual extraction of damage features. CNN replaces the multiplication of general neural networks with convolution operations and uses multiple convolution kernels to perform convolution processing separately to extract different types of features. The CNN structure contains multiple convolutional layers and pooling layers to extract more complete feature information. The applicability of the network architecture is verified in the processing of time series and energy features in this paper. At present, CNN has been widely used in image recognition, speech recognition, text processing, bearing fault diagnosis, and crack detection [24–28] and can realize damage detection of large engineering structures.

In this paper, a neural network for steel beam damage detection is proposed. Firstly, the numerical simulation data are used for pretraining, and then, the experimental data are input into the trained neural network to verify the prediction effect of CNN on damage. Compared with the traditional machine learning algorithm SVM, the recognition accuracy is verified. The related code data are shown in Supplementary Materials.

2. Basic Theories and Methods

2.1. Wavelet Packet Decomposition

As opposed to the physical quantities directly related to the characteristics of the structure, such as stiffness, natural frequency, and damping, energy plays a significant role in signal processing, which is only the energy of signal analysis angle and does not have the significance of energy in the real physical world. After structural damage, the acceleration signal will be partially disturbed, resulting in the energy change of each frequency band after the wavelet packet decomposition. The energy change in the frequency band corresponds to the damage done to the structure, offering a better reflection of the damage information. With the index of energy, a set of new feature vectors can be obtained by rearranging the frequency band information, so as to describe the damage scenarios of the structure more clearly. The process of WPD of the frequency band energy and the three-layer WPD tree in Figure 1 is as follows [29].
After the signal is decomposed by the j-layer wavelet packet, the initial signal \( f(t) \) is expressed as:

\[
f(t) = \sum_{i}^{2j} f_{j}^i(t)
\]  

(1)

Its linear combination form can be expressed as:

\[
f_{j}^i(t) = \sum_{-\infty}^{+\infty} c_{j,k}^i(t) u_{j,k}^i(t)
\]  

(2)

The wavelet packet coefficient can be expressed as:

\[
c_{j,k}^i = \int_{R} f(t) u_{j,k}^i(t) dt
\]  

(3)

where \( u_{j,k}^i(t) \) and \( c_{j,k}^i(t) \) denote the wavelet packet function and the wavelet packet coefficient, respectively. When \( m \neq n \), the orthogonality condition of the wavelet packet function can be expressed as:

\[
u_{j,k}^m(t) u_{j,k}^n(t) = 0
\]  

(4)

The formula for total energy is:

\[
E_j = \int_{R} f^2(t) dt = \sum_{m}^{2j} \sum_{n}^{2j} \int_{R} f_{j}^m(t) f_{j}^n(t) dt
\]  

(5)

According to the orthogonality condition of the wavelet function, we can obtain:

\[
E_j = \sum_{i=1}^{2j} E_j^i
\]  

(6)

The energy distributed over the i-band of the j-layer can be expressed as:

\[
E_j^i = \int_{R} f_{j}^2(t) dt
\]  

(7)

2.2. Selection of Wavelet Basis Function

After WPD, it is necessary to select the wavelet basis function. There are many kinds of wavelet basis functions—Morlet wavelet, Meyer wavelet, Symlets wavelet, Mexican Hat wavelet, Harr wavelet, and Daubechies wavelet are commonly used [30]. When different wavelet basis functions are used to process signals, the effects are different [10]. In this manuscript, Daubechies wavelet is selected as the basis function constructed by the wavelet packet energy’s characteristics, because not only does it has orthogonality and locality in the time domain but it can also make signal reconstruction smoother during digital analysis.
signal analysis. After selecting the wavelet basis function, the decomposition level of the wavelet is determined via its relation to the size of its cost function—the decomposition level with the smallest cost function value should be selected to make the calculation more efficient. Tables 1 and 2 show that the Daubechies15 wavelet with ten decomposition layers is selected for the numerical simulations and dynamic tests of beam structure under different damage scenarios. The cost function $I^p$ is expressed as follows:

$$I^p = S_L(E_j) = \sum_j |E_{j,i}|^p$$  \hspace{1cm} (8)

### Table 1. Cost function values for decomposing Daubechies with level 4.

<table>
<thead>
<tr>
<th>Wavelet Order</th>
<th>$S_L(E_j)$ ($p = 1.5$)</th>
<th>Wavelet Order</th>
<th>$S_L(E_j)$ ($p = 1.5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>532.7851</td>
<td>16</td>
<td>493.4817</td>
</tr>
<tr>
<td>12</td>
<td>512.3428</td>
<td>17</td>
<td>494.3682</td>
</tr>
<tr>
<td>13</td>
<td>480.7129</td>
<td>18</td>
<td>498.6675</td>
</tr>
<tr>
<td>14</td>
<td>554.1102</td>
<td>19</td>
<td>496.5085</td>
</tr>
<tr>
<td>15</td>
<td>479.8986</td>
<td>20</td>
<td>518.0319</td>
</tr>
</tbody>
</table>

### Table 2. Cost function values and computational time of db15 wavelet decomposition 10 layers.

<table>
<thead>
<tr>
<th>Decomposition Level</th>
<th>$S_L(E_j)$ ($p = 1.5$)</th>
<th>Computational Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>951.6591</td>
<td>1.496 s</td>
</tr>
<tr>
<td>2</td>
<td>891.3953</td>
<td>1.417 s</td>
</tr>
<tr>
<td>3</td>
<td>623.9734</td>
<td>1.447 s</td>
</tr>
<tr>
<td>4</td>
<td>479.8986</td>
<td>1.479 s</td>
</tr>
<tr>
<td>5</td>
<td>355.8308</td>
<td>1.531 s</td>
</tr>
<tr>
<td>6</td>
<td>252.9848</td>
<td>1.475 s</td>
</tr>
<tr>
<td>7</td>
<td>192.7643</td>
<td>1.480 s</td>
</tr>
<tr>
<td>8</td>
<td>144.2826</td>
<td>2.314 s</td>
</tr>
<tr>
<td>9</td>
<td>111.2022</td>
<td>6.054 s</td>
</tr>
<tr>
<td>10</td>
<td>87.3245</td>
<td>2.127 s</td>
</tr>
</tbody>
</table>

### 2.3. Construction of Damage Index

In order to better understand the damage to the structure, the ratio of its damage frequency band energy to the sum of all frequency band energies is defined as follows:

$$R_p = \frac{E_{j,p}}{\sum_{i=0}^{2^j-1} E_{j,i}} \quad (p = 1, 2, \ldots, m)$$  \hspace{1cm} (9)

The damage state of the structure is characterized by the change in the energy ratio of before to after the damage, and the energy ratio variation of the p band can be expressed as:

$$ERV_p = |R_{bp} - R_{ap}|$$  \hspace{1cm} (10)

In the above expression, $R_{bp}, R_{ap}$ represent the energy ratio of the p frequency band before and after structural damage, respectively.

### 2.4. Convolutional Neural Network

In this paper, the CNN in MATLAB deep learning toolbox is used to realize data pre-training and testing. The typical CNN contains at least a convolutional layer, a pooling layer, a fully connected layer, and a softmax output layer.

Figure 2 shows the network architecture of 1D-CNN structure. The 1D-CNN in this paper is composed of an input layer, a convolutional layer (kernel size is 3, number is 5, stride is 1), a max pooling layer (kernel size is 3 in S1 and S4, 2 in S2 and S3, stride
is 2), a batch normalization layer, a rectified linear unit (ReLU) layer, a flatten layer, a fully connected (FC) layer, and softmax output layer. Due to the data type, the prediction category is targeted to specific structural types, so the size of the input and output layer is fixed, and the accuracy of the final classification of CNN is directly related to the internal parameter setting. The CNN network architecture also can be used in the fields of bearing damage in mechanical structures and loosening of bolt structures.

Figure 2 shows the network architecture of 1D-CNN structure. The 1D-CNN in this paper is composed of an input layer, a convolutional layer (kernel size is 3, number is 5, stride is 1), a max pooling layer (kernel size is 3 in S1 and S4, 2 in S2 and S3, stride is 2), a batch normalization layer, a rectified linear unit (ReLU) layer, a flatten layer, a fully connected (FC) layer, and softmax output layer. Due to the data type, the prediction category is targeted to specific structural types, so the size of the input and output layer is fixed, and the accuracy of the final classification of CNN is directly related to the internal parameter setting. The CNN network architecture also can be used in the fields of bearing damage in mechanical structures and loosening of bolt structures.

2.4.1. One-Dimensional Convolutional Layer

The convolutional principle of one-dimensional convolutional layer is to multiply the convolutional kernel and internal elements one by one according to the stride of the input array. The following Formula (11) shows that the convolutional sequence of this process is multiplying one by one from left to right along the time series, from top to bottom, and then adding the product to the activation function to obtain the output value. The calculation formula of the characteristic diagram size of each output is shown in Formula (12). This process of convolutional kernel is shown in Figure 3.

\[
Y(m) = \sum_{x=1}^{k} K(m)S(m + x) \quad (11)
\]

\[
O = \frac{M - F + 2P}{S} + 1 \quad (12)
\]

where \(O\) represents the output signal, \(M\) represents the input signal, \(F\) represents the convolutional kernel size, \(P\) represents the number of zero fillings, and \(S\) represents the stride. The purpose of setting filling in CNN is mainly to extract the edge features without deviation, and the stride is set to reduce the number of input parameters and achieve the purpose of reducing the amount of calculation.
2.4.2. Batch Normalization Layer

In the training process of the model, batch normalization uses the mean and standard deviation of small batches to continuously adjust the intermediate output of the neural network, so that the intermediate output value of the entire neural network in each layer is more stable. The implementation process of batch normalization layer is as follows: Firstly, the mean and variance of small batch $B$ are calculated, and then, $x$ is standardized by dividing the square of elements. Here, the scale parameter $\gamma$ and the shift parameter $\beta$ are introduced, and finally, the output $y$ of batch normalization is obtained. The whole calculation process is composed of Formulas (13)–(16).

$$\mu_B = \frac{1}{m} \sum_{i=1}^{m} x^{(i)}$$  \hspace{1cm} (13)

$$\sigma^2_B = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)} - \mu_B)^2$$  \hspace{1cm} (14)

$$\tilde{x}^{(i)} = \frac{x^{(i)} - \mu_B}{\sqrt{\sigma^2_B + \epsilon}}$$  \hspace{1cm} (15)

$$y^{(i)} = \gamma \odot \tilde{x}^{(i)} + \beta$$  \hspace{1cm} (16)

2.4.3. Max Pooling Layer

The pooling layer in CNN can not only extract the feature information of relevant regions but also reduce the number of parameters, realize data dimension reduction, prevent model overfitting, and finally, improve the computational efficiency of neural network. The pooling layer generally has average pooling and max pooling. In this paper, the max pooling layer is selected for downsampling, and its implementation process is shown in Figure 4.

![Figure 4. One-dimensional max pooling process.](image)

2.4.4. Fully Connected Layer

The FC layer is the most common layer in the neural network. It connects a row of neurons one by one through the way of connecting the head and tail. In the FC layer, we need to specify the number of neurons and the activation function. After continuous
convolution and pooling of data, the original data set is mapped to the feature space, and then, several FC layers are cross-integrated to extract better signal features. The activation function of the FC layer uses the ReLU function, whose expression and the output of the FC layer are expressed as:

\[ R(x) = \begin{cases} 
0, & x \leq 0 \\
 x, & x > 0 
\end{cases} \]  \hspace{1cm} (17)

\[ R = f(\sum w \otimes x + b) \]  \hspace{1cm} (18)

2.4.5. Softmax Output Layer

After the FC layer is the softmax layer. The softmax layer of CNN mainly plays the role of classification prediction, which is used to display the final recognition accuracy. where \( a_k \) represents the kth input signal in the output layer and \( \exp(a_k) \) represents the exponential function of \( a_k \). Therefore, the output expression of the kth neuron is:

\[ y_k = \frac{\exp(a_k)}{\sum_{i=1}^{n} \exp(a_i)} \]  \hspace{1cm} (19)

3. Numerical Validation and Evaluation

3.1. Finite Element Model

The finite element model of a simple supported steel beam was established in ABAQUS, as shown in Figure 5. The basic parameters of the model are shown in Table 3. The beam along the span was divided into 10 equal elements, a total of 11 nodes, and the unit length was 0.1 m. In order to simulate the influence of quantitative impact load on a simple supported steel beam, the impact load shown in Figure 5 is applied to the impact time history, and the acceleration response signal of each sensor node under the impact load is calculated by the instantaneous modal dynamic analysis method. The excitation time is 0.001 s, the sampling time and interval sampling time of the whole structure is 0.1 s and 0.0005 s, the sampling frequency is 2000 Hz, and the structural damping ratio is 0.05. The element type of simply supported beam using C3D8R (an 8-node linear brick, reduced integration, hourglass control).

![Figure 5](image)

**Figure 5.** (a) Finite element model of a simple supported steel beam. (b) Shock time history and excitation position.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>1 m</td>
<td>density</td>
<td>7850 kg/m³</td>
</tr>
<tr>
<td>height</td>
<td>0.1 m</td>
<td>poisson’s ratio</td>
<td>0.3</td>
</tr>
<tr>
<td>width</td>
<td>0.068 m</td>
<td>force</td>
<td>100 N</td>
</tr>
<tr>
<td>waist thickness</td>
<td>0.0045 m</td>
<td>damping fraction</td>
<td>0.05</td>
</tr>
<tr>
<td>elastic modulus</td>
<td>210 GPa</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>
3.2. Numerical Simulation

In order to compare the effects of single damage and multiple damages on the positioning accuracy of different damage locations, 40 different damage scenarios are selected, as shown in Table 4. To simulate the damage in the form of reduction element stiffness, apply the impact load to the fourth node of the simple supported steel beam, and the acceleration signal of the steel beam before and after damage collected by sensor 3 is taken as an example, as shown in Figure 6. Assuming that the damage degree of simple supported steel beam is 1%, 2%, 5%, 10%, and 15% (The elastic modulus of the material changes to the original 0.99, 0.98, 0.95, 0.90, and 0.85 times), the first 10 order frequencies of the second element under five damage degrees are shown in Table 5. It can be seen from the figure that the difference between the signal curves of the intact scenarios and the damage scenarios is small. Therefore, in order to better observe the damage state of the structure, the WPD of the signal is needed to make it easier to distinguish different damage scenarios.

Table 4. Damage scenarios of simple supported steel beam.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Scenario Number</th>
<th>Elastic Modulus Under Different Damage Degrees</th>
<th>Damage Element</th>
<th>Location Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>intact</td>
<td>S0</td>
<td>E</td>
<td>none</td>
<td>intact</td>
</tr>
<tr>
<td></td>
<td>S1</td>
<td>0.85 E</td>
<td>1</td>
<td>left end</td>
</tr>
<tr>
<td></td>
<td>S2</td>
<td>0.85 E</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S3</td>
<td>0.85 E</td>
<td>5</td>
<td>middle span</td>
</tr>
<tr>
<td></td>
<td>S4</td>
<td>0.85 E</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S5</td>
<td>0.85 E</td>
<td>9</td>
<td>right end</td>
</tr>
<tr>
<td></td>
<td>S6</td>
<td>0.85 E</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>single</td>
<td>S7</td>
<td>0.85 E</td>
<td>2, 8</td>
<td>asymmetry</td>
</tr>
<tr>
<td></td>
<td>S8</td>
<td>0.85 E</td>
<td>5, 6</td>
<td>symmetry</td>
</tr>
</tbody>
</table>

Table 5. The first ten order frequencies of the intact and damaged scenarios.

<table>
<thead>
<tr>
<th>Mode No</th>
<th>Intact F (Rad/Time)</th>
<th>15% F (Rad/Time)</th>
<th>10% F (Rad/Time)</th>
<th>5% F (Rad/Time)</th>
<th>2% F (Rad/Time)</th>
<th>1% F (Rad/Time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>175.39</td>
<td>174.25</td>
<td>174.66</td>
<td>175.04</td>
<td>175.25</td>
<td>175.32</td>
</tr>
<tr>
<td>2</td>
<td>225.74</td>
<td>224.24</td>
<td>224.77</td>
<td>225.27</td>
<td>225.55</td>
<td>225.64</td>
</tr>
<tr>
<td>3</td>
<td>440.97</td>
<td>437.46</td>
<td>438.73</td>
<td>439.90</td>
<td>440.55</td>
<td>440.76</td>
</tr>
<tr>
<td>4</td>
<td>523.66</td>
<td>512.96</td>
<td>522.58</td>
<td>523.14</td>
<td>523.46</td>
<td>523.56</td>
</tr>
<tr>
<td>5</td>
<td>656.09</td>
<td>654.04</td>
<td>654.78</td>
<td>655.46</td>
<td>655.84</td>
<td>655.97</td>
</tr>
<tr>
<td>6</td>
<td>762.93</td>
<td>756.45</td>
<td>759.19</td>
<td>761.30</td>
<td>762.33</td>
<td>762.64</td>
</tr>
<tr>
<td>7</td>
<td>790.66</td>
<td>783.06</td>
<td>785.44</td>
<td>788.01</td>
<td>789.59</td>
<td>790.13</td>
</tr>
<tr>
<td>8</td>
<td>860.51</td>
<td>856.58</td>
<td>857.92</td>
<td>859.23</td>
<td>860.00</td>
<td>860.26</td>
</tr>
<tr>
<td>9</td>
<td>925.61</td>
<td>917.54</td>
<td>920.29</td>
<td>922.98</td>
<td>924.56</td>
<td>925.09</td>
</tr>
<tr>
<td>10</td>
<td>1048.70</td>
<td>1040.10</td>
<td>1043.30</td>
<td>1046.10</td>
<td>1047.70</td>
<td>1048.20</td>
</tr>
</tbody>
</table>

Therefore, the concept of wavelet packet mentioned above is introduced, which can decompose low-frequency signals and high-frequency signals at the same time. In the process of signal feature extraction, in order to find more sensitive feature indicators for structural damage, so that the subsequent CNN can be more efficient in feature extraction. In this paper, the selected db15 wavelet is used for ten-layer WPD, the signal is decomposed into 1024 energy bands, and the first 201 energy frequency bands with obvious characteristics are extracted. Combined with the characteristics of energy partial suppression or enhancement before and after structural damage, the ERV of wavelet packet band is used as the damage index.
Figure 6. Acceleration response signals of intact structure and damaged structure. (a) Working scenarios S0 and S1. (b) Working scenarios S0 and S2. (c) Working scenarios S0 and S3. (d) Working scenarios S0 and S4. (e) Working scenarios S0 and S5. (f) Working scenarios S0 and S6. (g) Working scenarios S0 and S7. (h) Working scenarios S0 and S8.
3.3. Results and Analysis

Figure 7 shows the ERV of the eight types of damage scenarios under the damage degree of 15%. It can be seen from the figure that the structural damage will have a certain impact on the ERV value. Taking the second unit damage in a single damage scenario as an example, the wavelet packet frequency band ERV varies greatly at the positions of first and second units, which can preliminarily judge the approximate location of the damage. Compared with other units, the energy released is also more. Similarly, when the damage occurs at other element locations, the damage index of the nearby element changes significantly, but when the damage occurs at the first and tenth unit positions at the support, the change of ERV of the mid-span unit is also severe and unstable.

![Figure 7. Two-dimensional ERV diagrams of different working scenarios. (a) ERV value of working scenario S1. (b) ERV value of working scenario S2. (c) ERV value of working scenario S3. (d) ERV value of working scenario S4. (e) ERV value of working scenario S5. (f) ERV value of working scenario S6. (g) ERV value of working scenario S7. (h) ERV value of working scenario S8.](image)
Therefore, taking the acceleration signal as the original data, the ERV constructed by wavelet packet frequency band energy method can only preliminarily identify the approximate location of the structural damage. In order to obtain more accurate identification results, it is necessary to combine the neural network for further judgment.

4. Comparison of CNN and SVM Training Effect

4.1. Damage Identification by CNN

The acceleration signal of the node in the Z direction of the finite element model of the steel beam is used as the original data, and the damage data set is obtained by WPD. Nine acceleration sensors are arranged on the beam, and six sets of data are collected at each sensor position. Then, the sample sets of eight types of damage states under five damage degrees of 1%, 2%, 5%, 10%, and 15% are $8 \times 6 \times 9 \times 5 = 2160$, and each damage state corresponds to 270 data. Among the 2160 samples, 1800 of the data are selected as training samples, and 360 of the data are selected as validation samples.

The training process of CNN is shown in Figure 8 below. The upper and lower layouts of the training process represent the training accuracy and loss function, respectively. The blue and red lines represent the training set, and the black dotted lines represent the verification set (the verification set is extracted from the original training set). The test set is input into the trained neural network model to obtain the recognition result of the damage scenario. The configuration and super parameter settings for the neural network training are shown in Tables 6 and 7. From the Figure 8, it can be seen that the accuracy of damage identification based on acceleration signal increases rapidly from about 20% to about 85%, and then, the change tends to be gentle, and the fluctuation range is small. The final accuracy rate can reach 95.14%.

![Figure 8. The training process of the CNN.](image)
Table 6. The configuration of the CNN for simple supported steel beam database.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Kernel Num. &amp; Size</th>
<th>Padding</th>
<th>Stride</th>
<th>Activation</th>
</tr>
</thead>
<tbody>
<tr>
<td>input</td>
<td>none; none</td>
<td>none</td>
<td>[1, 1]</td>
<td>none</td>
</tr>
<tr>
<td>convolution (C1)</td>
<td>6; [1, 3]</td>
<td>1</td>
<td>[1, 1]</td>
<td>relu</td>
</tr>
<tr>
<td>max pooling (S1)</td>
<td>6; [1, 3]</td>
<td>0</td>
<td>[1, 2]</td>
<td>none</td>
</tr>
<tr>
<td>convolution (C2)</td>
<td>12; [1, 3]</td>
<td>1</td>
<td>[1, 1]</td>
<td>relu</td>
</tr>
<tr>
<td>max pooling (S2)</td>
<td>12; [1, 2]</td>
<td>0</td>
<td>[1, 2]</td>
<td>none</td>
</tr>
<tr>
<td>convolution (C3)</td>
<td>24; [1, 3]</td>
<td>1</td>
<td>[1, 1]</td>
<td>relu</td>
</tr>
<tr>
<td>max pooling (S3)</td>
<td>24; [1, 2]</td>
<td>0</td>
<td>[1, 2]</td>
<td>none</td>
</tr>
<tr>
<td>convolution (C4)</td>
<td>48; [1, 3]</td>
<td>1</td>
<td>[1, 1]</td>
<td>relu</td>
</tr>
<tr>
<td>max pooling (S4)</td>
<td>48; [1, 3]</td>
<td>0</td>
<td>[1, 2]</td>
<td>none</td>
</tr>
<tr>
<td>convolution (C5)</td>
<td>96; [1, 3]</td>
<td>1</td>
<td>[1, 1]</td>
<td>relu</td>
</tr>
<tr>
<td>fully connected</td>
<td>none; none</td>
<td>none</td>
<td>none</td>
<td>relu</td>
</tr>
<tr>
<td>softmax</td>
<td>none; none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>output</td>
<td>none; none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>

Table 7. Training options Adam of the CNN for simple supported steel beam database.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial learn rate</td>
<td>0.001</td>
<td>verbose frequency</td>
<td>50</td>
</tr>
<tr>
<td>gradient threshold</td>
<td>1</td>
<td>shuffle</td>
<td>once</td>
</tr>
<tr>
<td>max epochs</td>
<td>200</td>
<td>execution environment</td>
<td>cpu</td>
</tr>
<tr>
<td>mini batch size</td>
<td>400</td>
<td>verbose</td>
<td>false</td>
</tr>
</tbody>
</table>

In order to obtain the final recognition accuracy of each type of damage scene, it is necessary to train and test each type of damage condition separately. The results are shown in Figure 9. The confusion matrix plots the prediction accuracy of eight damage locations and the misjudgment rate of the remaining locations.

Figure 9. Accuracy of different damage scenarios by the CNN.

For the left-end unit damage data set, 77.78% of the data in the first unit are identified correctly, 4.44% of the data are identified as ninth unit damage, 4.44% of the data are identified as tenth unit damage, 1.33% of the data are identified as asymmetric unit damage, 81.11% of the data in the second unit are identified correctly, 4.44% of the data are identified as the first unit damage, 1.11% of the data are identified as the sixth unit damage, 2.22% of the data are identified as the ninth unit damage, 7.78% of the data are identified as the
tenth unit damage, and 3.33% of the data are identified as the asymmetric unit damage. Similarly, the data identification of the mid-span unit, the right-end unit, the asymmetric unit, and the symmetric unit is shown in Figure 9, where the data identification accuracy of the asymmetric unit is as high as 93.33%.

4.2. Damage Identification by SVM

The core idea of SVM is to transform the maximum separation hyperplane problem into a constrained optimization problem. For linear classification, it is also necessary to use the Lagrange multiplier method to transform the objective function into the Lagrange function. By deriving the weight \( w \) and the bias \( b \) and then bringing back the objective function, the new objective function can be obtained as follows:

\[
\begin{align*}
\min_{\alpha} & \quad \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N} \alpha_i \\
\text{s.t.} & \quad \sum_{i=1}^{N} \alpha_i y_i = 0; \quad \alpha_i \geq 0, \quad i = 1, 2, \ldots, N
\end{align*}
\]  

(20)

In order to compare with the results of CNN, SVM is used to pretrain the numerical simulation data. Similarly, 1800 of the data is selected as the training set and 360 as the test set to verify the results and check the damage of the structure. Finally, it is found that the accuracy of data test using SVM is lower than that of CNN, and the accuracy is only 44.72% (161/360). The recognition effect of single damage is better, while that of multiple damage is significantly reduced. The main reason is the poor grid performance of SVM and the slow processing speed of massive data.

The classification results of predicted and measured types are shown in Figure 10, classification accuracy as shown in Table 8, • and ○ represent the actual type and forecast type, respectively, and the upper and lower layers represent the samples divided into two categories. When the two overlap, it indicates that the SVM correctly classifies the two types of samples. When the two are separated, it indicates that the SVM classifies the two types of samples incorrectly, and this method can thus intuitively show the accuracy of the determined actual type and forecast type.

![Figure 10. The process of classification using SVM.](image-url)
### Table 8. Accuracy of different damage scenarios by SVM.

<table>
<thead>
<tr>
<th>Damage Scenario</th>
<th>True Type</th>
<th>Predicted Type</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>2, 8</td>
<td>2, 8</td>
<td>44.72%</td>
</tr>
<tr>
<td>S8</td>
<td>5, 6</td>
<td>5, 6</td>
<td></td>
</tr>
</tbody>
</table>

5. Experimental Validation

5.1. Experimental Model

In order to verify the feasibility of numerical simulation, the experimental platform is set up as Figure 11. The size, material parameters, and physical parameters of the experimental model are the same as those of the finite element model. The simple supported steel-beam used in this paper is 1 m in length, 0.1 m in height, 0.068 m in width, and 0.0045 m in waist thickness. The two sides of the beam are fixed by the masonry cement bearing with good stiffness to limit the horizontal displacement of the structure. The acceleration sensor is arranged on the beam to simulate the damage scenarios of single damage scenario S2 and multiple damage scenario S7.

![Figure 11. Process of experimental arrangement: (a) Laptop. (b) Simple supported steel beam. (c) DH5927N dynamic signal test and analysis system. (d) Brick cement bearing. (e) Piezoelectric acceleration sensor.](image)

The notch damage simulation is made at the second unit, the second unit and eighth unit, respectively, as shown in Figure 12. The notch length is 0.05 m, the width is 0.068 m, the height is 0.026 m, and the damage area of the unit is 0.0047 m².

5.2. Experimental Facilities

The experimental instruments are the DH5927N dynamic signal test and analysis system, DHDAS dynamic signal acquisition and analysis system, and laptop. The components required for the experiment include rubber hammer, piezoelectric acceleration sensor with IA202E model, and transmission data line. The experimental device is shown in Figure 13.
5.2. ...ing scenarios, namely, working scenarios S0, S2, and S7, are selected for the steel-beam.

![Figure 12. The notch position and size diagram.](image)

5.3. Experimental Results and Analysis

When the excitation hammer is used to obtain the vibration signal, the measuring points at the junction of the third unit and the fourth unit are selected, and the data line is connected with the AI1-01 channel. Then, the channel parameters are set: the measurement value is acceleration, the measurement type is voltage measurement, the sensitivity is set to 1000 mv/g, the measurement range is 0.2, the input mode is SIN_DC, the sampling frequency is set to 2000 Hz, and the sampling time is 0.1 s. Before collecting data, each channel port needs to be balanced to ensure that the signal amplitude changes from zero.

![Figure 13. Experimental facilities: (a) Piezoelectric acceleration sensor. (b) Rubber hammer. (c) Laptop. (d) DH5927N dynamic signal test and analysis system.](image)

5.3. Experimental Results and Analysis

In order to verify the feasibility of the numerical model, a total of three types of working scenarios, namely, working scenarios S0, S2, and S7, are selected for the steel-beam. 54 sample data are collected for each type of working scenarios, and 162 test data sets are formed. The CNN model for training finite element data is used for further testing, and the accuracy of the experimental data is finally obtained. Firstly, the acceleration signals
under two damage working scenarios are obtained by vibration test, as shown in Figure 14, and then the acceleration signals are decomposed into each frequency band by wavelet packet. The ERV values obtained by comparison experiment and numerical simulation are shown in Figure 15. The results show that the matching degree between experimental data and numerical simulation data can reach about 95%. It can be seen from the figure that when the damage occurs at the position of the second unit, the ERV value at this position is significantly different from the ERV value of other units. When the damage occurs at the mid-span position of the second unit and the eighth unit, the change of the energy ratio of these two units can also be observed, and other units also show a small range of fluctuations.

![Figure 14. (a) Acceleration signals of S2. (b) Acceleration signals of S7.](image)

![Figure 15. (a) ERV value of working scenario S2. (b) ERV value of working scenario S7.](image)

After obtaining the energy ratio variation damage index of S2 and S7, the experimental data are used to verify the damage detection effect of CNN on real structures. Firstly, the grid is trained to obtain a sample database, and then the experimental data is used as a test sample for prediction. The final accuracy of the validation set is 91.67%. The network training process and the final detection results of single damage and multiple damage are shown in Figure 16.
6. Conclusions

In this paper, a one-dimensional CNN based on the ERV feature vector obtained by wavelet packet analysis is proposed to identify structural damage. The network structure can be applied to multiple scenarios, and different damage scenarios can be characterized by changing some properties of materials. Firstly, the neural network is trained by numerical simulation data, and then the experimental data are tested. The results show that CNN has achieved good recognition results for single damage and multiple damage, and this kind of CNN is applicable to various damage scenarios.

Compared with the recognition effect of CNN, it is found that when the SVM is used to pre-train the numerical simulation data, the accuracy rate of SVM can only reach 44.72% when the same training set and test set are selected. Although the SVM has good recognition effect on single damage, the recognition effect on multiple damage is not ideal. At the same time, the training process of SVM in dealing with massive samples is time-consuming and easy to cause data over-fitting. Therefore, the CNN selected in this paper can effectively deal with massive data. In order to improve the accuracy of the results, this paper selects wavelet packet for energy-related feature extraction as the training data and test data of neural network.

There are various types of damage in actual engineering structures, such as corrosion and cracks. Therefore, it is necessary to combine more advanced technologies for detection, such as ultrasonic and UAV technologies. In this manuscript, a simple supported beam is taken as an example to identify the location and degree of the damaged beam. In the future, this method can also be applied to the SHM of more complex continuous beams and multi-span beams. The damage to the structure can be accurately located by selecting the most appropriate principle of sensor arrangement and better characteristic indexes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app122010220/s1.

Author Contributions: Conceptualization, C.-S.W.; Data curation, W.R.; Formal analysis, C.-S.W.; Funding acquisition, C.-S.W. and Z.-Y.F.; Investigation, D.-B.Z.; Methodology, Y.-X.P.; Resources, J.-Q.Z.; Software, D.-B.Z.; Supervision, D.-B.Z.; Validation, J.-Q.Z., W.R. and Z.-Y.F.; Writing—original draft, C.-S.W.; Writing—review & editing, Y.-X.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Fund of National-local Joint Engineering Laboratory for Road Engineering and Disaster Prevention and Mitigation Technology in Mountainous Areas (No. 2018Road002), the Open Fund of Key Laboratory of Geological Hazards Mitigation for Moun-
tain Highway and Waterway, Chongqing Municipal Education Commission Chongqing Jiaotong University (kfxm2018-05), the Postdoctoral Fund Project of the Chongqing Natural Science Foundation (Cstc2019jcyj-bshx0113), and China Postdoctoral Science Foundation (2019M653346).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

22. Zhu, G.X. Research on Link Travel Time Prediction Based on Radial baSIS FUNCTION (RBF) Neural Network. Master’s Thesis, Chongqing University, Chongqing, China, 2004. [CrossRef]


