Deep learning based fatigue crack diagnosis of aircraft structures

L. Xu, S. F. Yuan*, J. Chen, Q. Bao

Research Center of Structural Health Monitoring and Prognosis, State Key Lab of Mechanics and Control of Mechanical Structures, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China.
Email: liangxu@nuaa.edu.cn; ysf@nuaa.edu.cn; cj1108@nuaa.edu.cn; baoqiao@nuaa.edu.cn

KEYWORDS: Convolutional neural network; Guided wave; Fatigue crack diagnosis; Aircraft structures;

ABSTRACT

Fatigue crack diagnosis is of great meaning for aircraft maintenance. However, reliable diagnosis in real engineering application still has challenges due to uncertainties which arise from sources such as environmental factors, loading and intrinsic material property changes. In recent years, deep learning (DL) methods have shown powerful recognition ability in many fields. Among them, the convolutional neural network (CNN) is able to enhance the pattern features and reach a better pattern recognition comparing with traditional artificial neuron networks, which has the potential to perform better crack recognition. In this paper, a guided wave (GW)-CNN based fatigue crack diagnosis method is proposed. GW features extracted from GW signals of different excitation-acquiring paths with different excitation frequencies are employed as the input of a trained CNN for fatigue crack diagnosis. An experiment verification was performed on a kind of attachment lug specimen, which is an important connecting component on aircraft structures. The results show that this proposed method is promising.

1. Introduction

Damage monitoring is a key issue for safety-critical systems such as aircraft, wind turbines, bridges and nuclear plants[1-2]. As one of the main damage types, fatigue crack usually occurs in many engineering structures. The guided wave (GW) based method[3-4] which has the advantages of sensitivity to small damage and large width of monitoring areas[5] has been widely studied for fatigue crack diagnosis[6-8]. However, reliable diagnosis in real engineering application still has challenges, because GW signals are highly affected by uncertainties which arise from sources such as environmental factors, loading and intrinsic material property changes[9]. It is difficult to perform fatigue crack diagnosis with features directly extracted from these GW signals.

In recent years, deep learning (DL) has shown its powerful recognition ability. As one of the important DL approaches, convolutional neural network (CNN) owns the ability of extracting feature representation layer by layer with convolution and pooling[10], which has the potential of amplifying features that are related with recognition. CNN has been tried in damage diagnosis of components including reactor metallic surfaces[11], bearing components[12], planetary gearbox[13] and etc. However, there are few researches on diagnosis of fatigue crack of aircraft structures with CNN.

In this paper, a GW-CNN based fatigue crack diagnosis method is proposed. After the GW features are obtained from multi-channels of GW signals, a designed CNN is applied to diagnose fatigue crack. The

* Corresponding author. S. F. Yuan, Email: ysf@nuaa.edu.cn.

Creative Commons CC-BY-NC licence https://creativecommons.org/licenses/by/4.0/
output corresponding to the highest output probability of the crack state is employed as the diagnosis result of the fatigue crack. The experiment verification was performed on a kind of attachment lug, which is an important connecting component for aircraft structures. The results show that this method can diagnose the fatigue crack of aircraft structure and is promising for further research.

2. The GW-CNN based fatigue crack diagnosis method

The proposed method is shown as Fig. 1. GW signals are excited and acquired by a piezoelectric transducers (PZTs) sensor array. GW features extracted from GW signals of different excitation-acquiring paths with different excitation frequencies are combined into an input sample, which is fed into the trained CNN for crack diagnosis. GW features are fused through convolution, pooling and full connecting layers, and recognized result is given by the softmax classifier, which is the last layer of the designed CNN.

2.1 GW feature extraction

Usually, GW features are extracted as damage indexes (DIs) to quantify signal changes causing by crack\textsuperscript{[14]}. Though DIs directly reflect the crack growth and can be used for crack diagnosis, signals are highly affected by uncertainties including loading, structure’s geometry, intrinsic material properties, wave behaviours, crack propagation paths and so on. Fig. 2 shows the typical GW signals obtained in the evaluation research in this paper.

![Figure 1. The proposed GW-CNN based fatigue crack diagnosis method](image)

**Figure 1.** The proposed GW-CNN based fatigue crack diagnosis method

Compared with traditional machine learning methods, the CNN is capable of fusing a large amount of data due to convolution and pooling procedure and its deeper network layers. In addition, the CNN owns the robustness to noise and generalization ability to new data. These characteristics allow CNN to fuse a large number of GW features of different channels with different frequencies. Hence, to include more information regarding the crack for a good recognition, this paper obtains GW signals from different excitation-acquisition channels with different excitation frequencies for one crack state.
Assuming that there are $m$ channels, $f$ excitation frequencies of GW signals, and $w$ wave packets are selected for every signal, and $d$ kinds of DIs are defined. There are totally $m \times f \times w \times d = n$ DIs from one crack state, these DIs can form an input sample for CNN to recognize.

### 2.2 CNN based fatigue crack diagnosis

CNN was proposed by Lecun\cite{15} in 1998. The basic structure of CNN is shown in Fig. 1, which mainly includes convolution layers, pooling layers and full-connected layers.

The convolution procedure in CNN is actually a mathematical interrelation progress of two variables. One-dimensional (1D) input $I$ is convolved by 1D variable $W$ ($W$ is called the convolution kernel) as shown in equation (1), where $i$ and $m$ are integers. The convolution layer usually has $k$ ($k=1,2,3,…$) convolution kernels as shown in figure 1.

$$F(i) = (I*W)(i) = \sum_{m} \sum_{n} I(i+m)W(m) \quad (1)$$

The convolution layer is characterized by local connections and shared weights \cite{15}. Compared with the fully connected network, these two characteristics can greatly reduce the number of parameters and the computation consumption. In addition, due to the existence of local connections, the local noise from the input sample only affects several neurons of the output instead of all the neurons. This increases the anti-noise capability of CNN\cite{16}.

The pooling layer is always followed after convolution layer. Pooling is similar to convolution except that it takes the maximum/average value as its output. This always allows pooling layer taking the typical feature to be its output. The pooling layer furtherly reduces the data volume, and makes the output own the invariability of translation, rotation and distortion\cite{16}.

Convolution and pooling make CNN owns more generalization ability to unknown environment. The generalization ability can help to deal with the uncertainties of input samples.

There is no difference between the full connected layers in CNN and traditional neural network. The softmax classifier has the same number of neurons as the desired number of classifications, and its output is the probability of different classifications. Generally, the class corresponding to the maximum output probability is the recognition result of the entire neural network.

As for training the designed CNN, crack states are divided into $n$ levels, which are seen as the desired output of CNN. The input sample obtained in each state corresponds to a crack level. The training process of the CNN is to minimize the cross-entropy cost function with optimization method like the Adam algorithm\cite{17}. Cross-entropy describes the distance between the network output and the desired output, which is shown as equation (2).

$$H(p, q) = - \sum_{x} (p(x) \log q(x) + (1 - p(x)) \log(1 - q(x))) \quad (2)$$

Where $p(x)$ denotes the probability of the desired output and $q(x)$ denotes the probability of the network output.

To increase the generalization ability of CNN, data enhancement, dropout\cite{18}, regularization L2 and early stop are applied in this paper.
3. Experimental verification

3.1 Fatigue test experiment of attachment lug

The attachment lug which is an important connecting component in aircraft structures is chosen for experimental verification. Fig. 3 illustrates the geometry and the sensors layout of attachment lug specimens. A 2mm notch was prefabricated on one side of the lug hole to control the crack growth initiation. These lugs are made of 5mm-thick LY12 aluminium. The fixture is connected with the lug and transmits the axial tension load with a dowel pin.

As shown in Fig. 4, a MTS810 electro-hydraulic servo tensile machine is used to apply the fatigue load. A sinusoidal load with a peak value of 18kN is chosen for the fatigue experiment, and the load frequency is chosen as 10Hz, the stress ratio \( R = \frac{F_{\text{min}}}{F_{\text{max}}} = 0.1 \). During the fatigue experiment, the multi-channel PZTs array scanning system developed by the authors’ group is employed to perform the active GW based monitoring \(^{[19]}\). Signals are collected every 2.5s within the crack length of 2 mm to 20mm.

![Figure 3. Geometry and sensor layout of the attachment lug](image)

![Figure 4. The fatigue experimental setup](image)

During the experiment, 3-cycle sine burst signals with 160kHz central frequency and 5-cycle sine burst signals with 160kHz central frequency are adopted as the excitation signals, ±70 V amplitude is chosen as the excitation signal, which is sampled with a sampling frequency of 50 MHz. As shown in Fig. 3, 3 channels of GW signals for each excitation frequency are obtained at the same crack state. For convenience, these excitation-acquisition channels are denoted as channel 2-1 160kHz, channel 3-1 160kHz, channel 5-4 160kHz, channel 2-1 190kHz, channel 3-1 190kHz, and channel 5-4 190kHz.

3.2 GW signals and features

Totally 7 specimens are tested, which labelled as T1~T7. For each specimen, a total of 3 channels with 2 excitation frequencies of GW signals are obtained at the same crack state. GW signals of channel 3-1
160kHz and channel 2-1 190kHz are shown in Fig. 5 and Fig. 6 respectively. Where the wave packet in red box is selected to extract GW features. These wave packets are S0 and A0 modes respectively, and their modes are relatively clear. They have less boundary reflection interference, which can better reflect the influence of cracks on GW signals.

Figure 5. Signals of channel 3-1 160kHz  
Figure 6. Signals of channel 2-1 190kHz

7 kinds of DIs (named DI1~DI7) are defined for feature extraction in this paper. The algorithms of these DIs are shown in table 1, where $H(t)$ and $D(t)$ are the baseline signal and the monitored signal respectively. $t_1$, $t_2$ are the start time and the end time of the selected wave packet for feature extraction.

$\omega$ denotes the signal frequency obtained by the Fourier transform.

**Table 1. GW based feature extraction algorithm**

<table>
<thead>
<tr>
<th>Name</th>
<th>Feature extraction algorithm</th>
<th>Serial number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross correlation</td>
<td>$DI_{1} = 1 - \sqrt{\frac{\int_{t_1}^{t_2} H(t)D(t)dt}{\int_{t_1}^{t_2} H^2(t)dt}}$</td>
<td>DI1</td>
</tr>
<tr>
<td>Spatial phase difference</td>
<td>$DI(H,D) = \int_{t_1}^{t_2} (\tilde{D}(t) - \alpha H(t))^2 dt \cdot \tilde{D}(t) = \frac{D(t)}{\int_{t_1}^{t_2} D^2(t)dt}, \alpha = \frac{\int_{t_1}^{t_2} \tilde{D}(t)H(t)dt}{\int_{t_1}^{t_2} H^2(t)dt}$</td>
<td>DI2</td>
</tr>
<tr>
<td>Spectrum loss</td>
<td>$DI = \frac{\int_{\omega_1}^{\omega_2}</td>
<td>H(\omega) - D(\omega)</td>
</tr>
<tr>
<td>Central spectrum loss</td>
<td>$DI = \frac{a(\omega) - b(\omega)}{a(\omega)}, a(\omega) = \max(H(\omega)), b(\omega) = \max(D(\omega))$</td>
<td>DI4</td>
</tr>
<tr>
<td>Differential curve energy</td>
<td>$DI = \sum_{n=2}^{N} \sum_{k=2}^{N} \frac{[d(t_n) - d(t_{n-1})]^2}{\sum_{n=2}^{N}[H(t_n) - H(t_{n-1})]^2}, d(n) = H(n) - D(n)$</td>
<td>DI5</td>
</tr>
<tr>
<td>Normalized Correlation Moment</td>
<td>$DI = \int_{-\infty}^{+\infty} T^*</td>
<td>r_{HD}(\tau)</td>
</tr>
<tr>
<td>Differential signal energy</td>
<td>$DI = \int_{-\infty}^{+\infty} (H(t) - D(t))^2dt \cdot H(t) = \frac{H(t)}{\sqrt{\int_{-\infty}^{+\infty} H^2(t)dt}}, \tilde{D}(t) = \frac{D(t)}{\sqrt{\int_{-\infty}^{+\infty} D^2(t)dt}}$</td>
<td>DI7</td>
</tr>
</tbody>
</table>

As for one signal, 7 DIs (DI1~DI7) are obtained from each wave packet. Fig. 7 shows the values of DI1 obtained by T1~T7 at channel 3-1 160kHz. It can be seen from the figure that the dispersion of DI is large, making it difficult to evaluate the crack directly.
3.3 Fatigue crack diagnosis with the CNN and GW features

The fatigue crack state in the attachment lug is divided into 11 crack levels according to the crack length interval defined in Table 2.

<table>
<thead>
<tr>
<th>Crack level</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>C9</th>
<th>C10</th>
<th>C11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crack length(mm)</td>
<td>0–1</td>
<td>1–2</td>
<td>2–3</td>
<td>3–5</td>
<td>5–7</td>
<td>7–9</td>
<td>9–11</td>
<td>11–13</td>
<td>13–15</td>
<td>15–17</td>
<td>17–20</td>
</tr>
</tbody>
</table>

For all the 7 specimens labelled as T1–T7, totally samples are obtained from the acquired GW signals. All the samples are labelled with corresponding crack level. In this verification of the proposed method, 70% of all the samples are selected for training the designed CNN, and the rest of samples are adopted for the test of the fatigue crack diagnosis. These training and test samples are randomly assigned.

Fig. 8 (a) shows the values of DI1 under different crack levels in the test data, which are extracted from channel 3-1 160kHz. It can be found that the values of DI1 under different crack levels overlap seriously. As shown in Fig. 8 (b), the values of DI1 from channel 2-1 190kHz also give the similar result. These overlap come from the uncertainties during the monitoring making it difficult to distinguish different crack levels with DIs directly.

The designed CNN which is shown in Fig. 9 is trained first with the training samples. When training the CNN, the hyper-parameters are given in Table 3. For convenience, convolution layers are denoted as Con1, Con2 and Con3, full-connected layers are denoted as Ful1, Ful2, Ful3 and Ful4, the dropout parameter is denoted as Kp.
Figure 9. The structure of the designed CNN

Table 3. The hyper-parameters of the designed CNN

<table>
<thead>
<tr>
<th>Hyper-parameters</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate $\alpha$</td>
<td>0.001</td>
</tr>
<tr>
<td>Mini-batch size</td>
<td>500</td>
</tr>
<tr>
<td>The activation function</td>
<td>ReLU function</td>
</tr>
<tr>
<td>Con1, Con2</td>
<td>Size: $7 \times 1$, channels:16, $K_p=0.4/0.45$</td>
</tr>
<tr>
<td>Con3</td>
<td>Size: $5 \times 1$, channels:32</td>
</tr>
<tr>
<td>Pooling layer</td>
<td>Size: $5 \times 1$, channels:32, $K_p=0.5$</td>
</tr>
<tr>
<td>Full1</td>
<td>Number of neurons:2048, $K_p=0.45$</td>
</tr>
<tr>
<td>Full2</td>
<td>Number of neurons:1024, $K_p=0.4$</td>
</tr>
<tr>
<td>Full3, Full4</td>
<td>Number of neurons:100, $K_p=0.6$</td>
</tr>
</tbody>
</table>

Fig. 10 (a) shows the value of the cross-entropy cost function evaluated with training samples and test samples. If the testing cross-entropy costs (testing costs) are higher than the training cross-entropy costs (training costs), it usually means overfitting. So, the training is usually stopped at the iteration where testing costs meet training costs.

In this paper, the training/testing accuracy rate which is used to measure the diagnostic performance of proposed method is defined as equation (3).

$$\text{Training/testing accuracy}=\frac{N}{T} \times 100\%$$  \hspace{1cm} (3)

Where $N$ represents the number of the network output equalling to the desired output, $T$ represents the total number of training/testing output.

As shown in Fig. 10, only one channel (channel 3-1 160kHz) of DI's are fused in the input sample. The testing costs are going to be higher than training costs at around 350 iterations, where the training accuracy rate is 88.45% and the testing accuracy rate is 71.01%. The training and testing results for other channels are shown in Table 4.

![Figure 10](image-url)
Furthermore, when DIs of 5 channels are fused in the input sample, the training and testing results are shown in Fig. 11. The testing costs are going to be higher than training costs at around 50 iterations, where the training accuracy rate is 98.98% and the testing accuracy rate is 94.92%.

![Figure 11. The training and testing results](image)

(a) The cross-entropy cost values with iterations  
(b) The accuracy with iterations

### Table 4. Comparison of classification accuracy under different fusion modes

<table>
<thead>
<tr>
<th>Input data channel</th>
<th>2-1 160kHz</th>
<th>3-1 160kHz</th>
<th>5-4 160kHz</th>
<th>2-1 190kHz</th>
<th>5-4 190kHz</th>
<th>All channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>The training accuracy after 500 training</td>
<td>85.32%</td>
<td>88.45%</td>
<td>86.45%</td>
<td>90.12%</td>
<td>89.25%</td>
<td>98.98%</td>
</tr>
<tr>
<td>Diagnostic accuracy</td>
<td>73.62%</td>
<td>71.01%</td>
<td>65.03%</td>
<td>84.74%</td>
<td>78.14%</td>
<td>94.92%</td>
</tr>
</tbody>
</table>

It can be seen from the results that when DIs of multiple channels are fused in the input sample, the testing diagnostic accuracy rate is obviously improved.

### 4. Conclusion

In this paper, a GW-CNN based fatigue crack diagnosis method is proposed. GW features are obtained from multi-channels of GW signals, a designed CNN is applied to diagnose fatigue crack. The fatigue experiments of attachment lug which is an important connecting component on aircraft structures is performed for verification. The results show that this proposed method is promising for reliable fatigue crack evaluation under uncertainty influence.

### Acknowledgement

This work was supported by the Key Program of National Natural Science Foundation of China [grant number 51635008]; National Natural Science Foundation of China [grant number 51575263]; Research Fund of State Key Laboratory of Mechanics and Control of Mechanical Structures (Nanjing University of Aeronautics and astronautics) [grant number MCMS-0517K01].

### References


