aerospace

MDPI

*Article*

# Random Dynamic Load Identification with Noise for Aircraft via Attention Based 1D-CNN

**Wenbo He [1,2], Xiaoqiang Zhang [3], Zhenyu Feng [1,2,*], Qiqi Leng [3], Bufeng Xu [2] and Xinmin Li [3]**

[1] Key Laboratory of Civil Aviation Aircraft Airworthiness Certification Technology, Civil Aviation University of China, Tianjin 300300, China
[2] School of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China
[3] School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China
* Correspondence: fzycauc@163.com

**Abstract:** Dynamic load identification plays an important role in the field of fault diagnosis and structural modification design for aircraft. In conventional dynamic load identification approaches, accurate structural modeling is usually needed, which is difficult to obtain for highly nonlinear or unknown structures. In this paper, a one-dimensional convolution neural network with multiple modules is proposed for random dynamic load identification of aircraft. Firstly, the convolution module is designed for temporal feature extraction. Secondly, the extracted features are linearly weighted based on the contributions to the final output. The contributions are learned in a data driven manner via the designed attention module. Lastly, the dynamic load of a certain time stamp is predicted from the learned and weighted features. The proposed model is trained and tested using the real data from a GARTEUR aircraft model. Extensive experimental results with qualitative and quantitative evaluations have demonstrated the identification performance with satisfactory accuracy of the proposed approach under different strengths of load noises.

**Keywords:** dynamic load identification; random dynamic loads; one-dimensional convolution neural network; attention mechanism; deep learning



**Citation:** He, W.; Zhang, X.; Feng, Z.; Leng, Q.; Xu, B.; Li, X. Random Dynamic Load Identification with Noise for Aircraft via Attention Based 1D-CNN. *Aerospace* **2023**, *10*, 16. https://doi.org/10.3390/aerospace10010016

Academic Editor: Andrea Da-Ronch

Received: 24 October 2022
Revised: 10 December 2022
Accepted: 20 December 2022
Published: 25 December 2022



## 1. Introduction

With the development of the aviation industry in recent years and the wide application of civil aircraft in public transportation, the research on the safety of aircraft operation has also received more and more attention [1–5]. Usually, the aircraft is working in complex mechanical environments and is subjected to different types of dynamic forces or loads, e.g., the dynamic loads from engine rotor, and the random dynamic loads from the airflow disturbance. It is important to analyze the influence of dynamic load on the aircraft body. The accurate measured external dynamic loads of the aircraft plays an important role in the overall design of the aircraft structure, the strength check and the environmental prediction. However, in practice, limited by factors such as shape design and sensor installation arrangements, it is difficult, if not impossible, to directly measure the external dynamic loads on the aircraft structures. In some cases, the vibration response at certain positions can be measured, and it can be utilized for identifying the external dynamic load information [6]. The load identification can also be used in fault diagnosis. For example, there might be excessive clearance of the aircraft actuator. As the actuator is usually connected with the frame of the aircraft structure, the load can be regarded as a load at some point on the aircraft structure. The identification of such loads would be useful for fault diagnosis and localization. However, in real situations, there could be multiple excessive clearances at different positions, the unknown loads would interference each other, resulting in the increase of the challenge of accurate load identification.

In early research of dynamic load identification, the direct inversion method is often adopted [7]. It builds an algebraic equation between the loads and the responses, and then estimates the dynamic loads from the response by inverting the coefficient matrix in the algebraic equation. While the direct inversion method is simple in principle and easy to apply in practice, it is often difficult to build the equation. For unknown or complicated structures, the problem is often ill-posed, which limits the accuracy of dynamic loads identification. Regularization-based methods improve the ill-posed problem of direct inversion. By adding reasonable boundary conditions to the ill-posed problem, the regularized optimal solution of the equation is sought. Jacquelin et al. [8] provides the L-curve and the general cross-validation (GCV) criterion for determining regularization parameters in truncated singular value decomposition (TSVD) or Tikhonov regularization approaches. Wang et al. [9] propose a regularization-based method with different regularization operators for identification tasks of complex structures. A Bayesian principle-based augmented Tikhonov regularization method is proposed in [10] to improve the load identification accuracy under severe ill-posed conditions. Bianchi et al. [11] proposes a fractional In addition to the TSVD and Tikhonov regularizations, the $L_1$ norm-based regularization methods for load identification [12–14] are also proposed. Li et al. [15] propose a $L_q$ norm-based regularization method, in which the $q$ can be learned from a posteriori knowledge.

For the above-mentioned conventional load identification approaches, an accurate dynamic model of the structure is also required. For complex structures, the establishment of an accurate structural dynamic model is often difficult. It can be seen that conventional methods rely too much on the transfer function of the structures. The measurement accuracy of the transfer function, the ill-posedness caused by the rank deficiency of transfer function matrix and the selection of regularization parameters would all affect the accuracy of the identification results.

With the recent rapid developments in artificial intelligence, neural networks (NNs) are now applied in the field of load identification. As a data-driven approach, NN-based methods avoid the need for accurate structural modeling, which could be suitable for load identification of unknown complex structures. In the literature, approaches with different types of NNs [16–21] are proposed to identify the static or dynamic load for different structures. Trivailo et al. [16] address the load identification problem of the empennage of a F/A-18 fighter aircraft via a shallow NN. Chen et al. [17] develop a neural network with two hidden layers to solve the impact load conditions for an elastoplastic hemispherical metal shell. Data generated from the finite element method are utilized for network training and testing. Wang et al. [18] propose to use a feed-forward NN to solve the nonlinear identification problem for a one-stage spur gearbox, an extended Kalman filter based algorithm is applied for NN parameters training. Zhou et al. [19] propose a long short-term memory (LSTM)-based recurrent NN model for impact load identification. Hossain et al. [20] compare the performance of radial basis function based NN and multilayer perceptron (MLP) for impact identification of a rectangular plate. A comparison between the NN and LSTM-based identification methods on different types of loads for aircraft is provided in [1].

The main problem of the above-mentioned feed-forward NN or the MLP is that the amount of parameters in the NN is relatively large. To reduce the number of parameters in the NN model, the convolutional neural network (CNN) can be applied. The CNN would process only part of the input data via convolution with a kernel. The parameters of CNN, or the weight of the kernel, are less in quantity compared to those of the full connect feed-forward NNs or the MLP, in which the entire input data are processed. In the literature, CNNs [22] are applied for fault diagnosis [23,24], damage detection [25] based on vibration data in both the time domain [23,25] and the frequency domain [24]. The problems of fault diagnosis or damage detection are to classify the vibration in time or frequency domain into a certain number of classes, e.g., different types of faults. As for the problem of load identification, the goal is to estimate the continuous, time-varying load signal based on the vibration data. Therefore, these problem are different. Xia et al. [26] applied a time

delay neural network (TDNN) for dynamic load identification for an aircraft rudder model. The TDNN can be regarded as a special CNN with one layer. Yang et al. [27] proposes to use a two-layer CNN and two-layer MLP to regress the dynamic loads. Yang et al. [28] utilize the hand-crafted wavelet features of signals as an input of a LSTM network and identify the dynamic on a three-degrees-of-freedom vibration system. Models in [26,27] do share some similarities with the proposed model. However, on one hand, Ref. [26] tends to use one layer but large kernels (size of 50) while multiple, small kernels (size of 2) in a multiple layer fashion are utilized in the proposed model. On the other hand, in the proposed model, we tend to further obtain the contributions for the CNN learned features, i.e., the outputs of the convolutions, to the final prediction via a designed attention mechanism and to fuse the output from small convolutions with the learned contributions. These would be the main differences of the proposed approach with methods in [26,27].

To address the load identification problem for unknown structures, especially for random dynamic loads identification problem on the aircraft structure with the existence of the interference of another noise load, in this paper, a novel one-dimensional convolution neural network (1D-CNN) with attention mechanism based approach is proposed. The network would make use of the convolution to extract high-level temporal features of the input multiple synchronized vibration response signals. Considering the fact that different features would contribute differently to the final load prediction, an attention module is designed to learn the contributions for the features from the training data. The dimensions of the input(s) and the output(s) of the proposed approach can be flexible to fit single input single out (SISO), multiple inputs single output (MISO) and multiple inputs multiple outputs (MIMO) load identification problems. Data obtained from a real self-build GARTEUR aircraft model is used for model parameters training and testing. Extensive experimental results show that the proposed approach can accurately identify the random dynamic loads in the presence of noises. Quantitative comparisons shows that the proposed approach outperforms conventional regularization-based approaches and shallow NN based-approaches.

The rest of the paper is structured as follows. In Section 2, we describe the problem of load identification, as well as the proposed deep learning model. Section 3 presents the details of the experimental results. We give the limitations and conclude the paper in Section 4.

## 2. Proposed Method

In this section, we describe the proposed approach for random dynamic load identification via attention-based 1D-CNN. The problem of load identification is first analyzed. The proposed 1D-CNN model is then described, as well as the implementation details for the network.

### *2.1. Load Identification Problem*

Considering a linear vibration system with $p$ dynamic loads, the corresponding acceleration response $\ddot{y}(t)$ in time domain at a certain measuring point can be calculated from the dynamic loads according to the principle of linear superposition,

$$\ddot{y}(t) = \sum_{i=1}^{p} \int_0^t h_i(t-\tau) f_i(\tau) d\tau, \tag{1}$$

where $h_i(\cdot)$ is the unknown response function for the $i$-th dynamic load at the measuring point, $f_i$ the $i$-th dynamic load. The measurements for acceleration obtained in real experiments are in discrete form. Equation (1) can be discretized by

$$\ddot{y}(t) = \sum_{i=1}^{p} \sum_{n=0}^{N_{load}} h_i(n\Delta t) f_i(t - n\Delta t), \tag{2}$$

where $\Delta t$ is a time interval for discretization, $N_{load} = \frac{t}{\Delta t}$ and $N_{load} \in \mathbb{N}$.

As for the problem of load identification, it is to estimate the dynamic load $f_i$ in Equation (2) in time domain from the acceleration measurements $\ddot{y}$ at different times. From Equation (2), it can be seen that the dynamic load at a certain time $t$, would affect the response $\ddot{y}$ of one measuring point at time $t$, $t + \Delta t$, $t + 2\Delta t$, $\cdots$. If the vibration system is not linear, the integral equation of Equation (1) would be invalid. The loads can be seen as a function of responses at time $t$, $t + \Delta t$, $t + 2\Delta t$, $\cdots$. More generally, if there are $K$ different measuring points in the system, the dynamic load $f_i(t)$ can be represented by a function $g(\cdot)$ of acceleration measurements from $K$ measuring points,

$$f_i(t) = g(\ddot{y}_1(t), \ddot{y}_1(t + \Delta t), \cdots, \ddot{y}_2(t), \ddot{y}_2(t + \Delta t), \cdots, \ddot{y}_K(t), \ddot{y}_K(t + \Delta t), \cdots), \tag{3}$$

where $\ddot{y}_k(t + n\Delta t), k \in [1, K], n \in [0, N]$ is the response of measuring point $k$ at time $t + n\Delta t$, respectively. The problem of load identification is therefore to estimate the value for $f_i(t)$ given the response measurements of $\ddot{y}_1(t), \ddot{y}_1(t + \Delta t), \cdots, \ddot{y}_2(t), \ddot{y}_2(t + \Delta t), \cdots$, $\ddot{y}_K(t), \ddot{y}_K(t + \Delta t), \cdots$.

### *2.2. The Proposed 1D-CNN-Based Model*

Usually, it is difficult to explicitly model the function of $g(\cdot)$ in Equation (3) for unknown or complicated structures. Therefore, in this paper, we would utilize the deep NN to model $g(\cdot)$ in an implicit and data-driven manner. More specifically, an attention-based 1D-CNN is proposed. Figure 1 shows the framework of the propose approach. The network consists of the convolution modules (blue dotted box in Figure 1), the attention modules (purple dotted box color in Figure 1), a prediction module for final load prediction. The network is designed in an end-to-end manner, the inputs of the network are the vibration signals from $K$ measuring points, and the outputs are the estimated dynamic load. Details of network modules are given below.

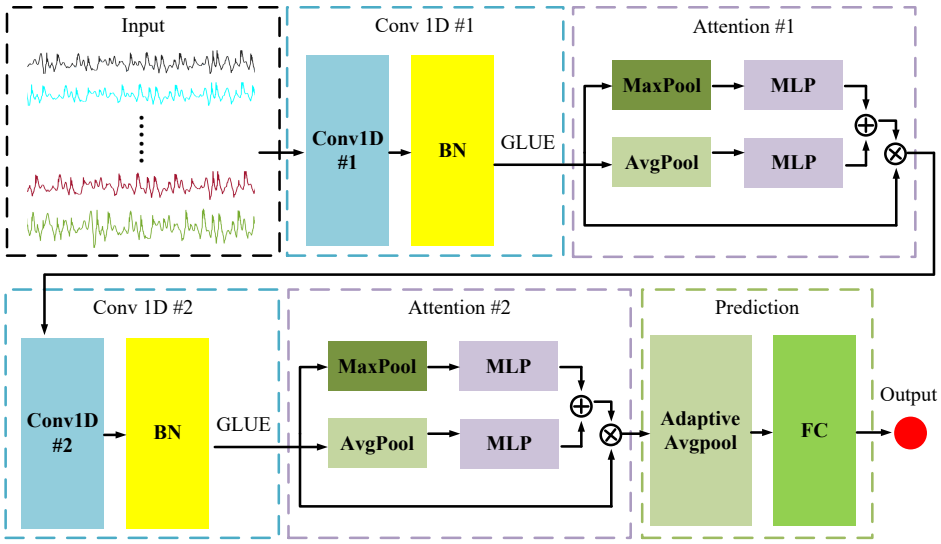



**Figure 1.** The network architecture of the propose approach. The inputs are $K$ measurements in the time domain. The convolution modules are utilized for feature extraction, while the attention module are used for feature weighting. The final prediction module would map the learned features to the dynamic load for a certain time stamp.

#### 2.2.1. Convolution Module

This module would utilize the convolution to extract temporal features of the inputs and generate feature maps for further processing. Considering the inputs of the network are signals of acceleration measurements from $K$ measurement points in the time domain, the 1D convolution is adopted, which is frequently used in signal processing

problems. The module includes a 1D convolution layer, a batch normalization layer and an activation function.

1D convolution layer. This layer would take the vibration signals (for the first module) or the feature maps (for the second module) as input. For a discrete input signal $f$, the output of the 1D convolution is given by

$$o(n) = \sum_{i=1}^{M} k(i) * f(n+i), \tag{4}$$

where $k$ is the convolution kernel, $M$ the size of $k$. The operator $*$ is the cross-correlation. In the network, there would be multiple corresponding kernels for a single 1D input. The number of the kernels is the filter number $N_{fn}$. The parameters of all kernels would be learned from data. Details for the input and output dimensions are discussed in Section 2.3. To enlarge the receptive field of the convolution, the dilated convolution, which is also called the convolution with holes [29–31], is used. The spacing between the values in a kernel is the dilation rate $d$.

Batch normalization. Considering the data of load identification is in small quantity, to reduce the possible over-fitting, the batch normalization (BN) is applied in this module. In this layer, the output of the dilated convolution layer would be normalized to a standard Gaussian distribution. Given the mean and variance estimated of the output $x$ from the previous layer, i.e., $mean(x)$ and $var(x)$, respectively, the output for the BN is given by

$$y = BN(x) = \frac{x - mean(x)}{\sqrt{var(x) + \epsilon}}, \tag{5}$$

where $\epsilon$ is a pre-defined small constant.

The Activation function. The output of the BN layer is then processed via the activation function. In the proposed approach, the Gaussian error linear unit (GELU) function [32] is used. Suppose $x$ is the input, the GELU is given by [32],

$$GELU(x) = x\Phi(x), \tag{6}$$

where $\Phi(x) = P(X \leq x), X \sim \mathcal{N}(0,1)$ is the standard Gaussian cumulative distribution function. The GELU is a modified rectified linear unit (ReLU) activation function with nonlinearity for the inputs, and show good performances in speech and vision tasks [32]. The outputs of the activation function are passed to the next module.

#### 2.2.2. Attention Module

The attention mechanism is a special designed module in the NN models to learn and compute the contribution of input data to output data. Inspired by the SE-Net [33], an attention module is designed in the proposed approach. Figure 2 shows the sketch of this module.

Suppose that the input for the attention module is a feature map with the dimension of $H \times C$. The feature maps can be regarded as $C$ different $H \times 1$ features (see Figure 2 left). The attention module tends to weigh the $C$ features using a learned weight, or, the contributions. The contributions of all $C$ features are learnt in a data-driven manner. First, the feature map $H \times C$ is down-sampled to $1 \times C$ via two pooling operation. In the "MaxPool" branch, the maximum value of each $H \times 1$ features is selected, while in the "AvgPool" branch, the mean value of each $H \times 1$ features is stored. Secondly, the two $1 \times C$ vector are transposed to $C \times 1$ and connected to two separate fully connected NNs, i.e., the MLPs. The MLP can be regraded as an implicit function. Here, the MLPs are used to map the input into two contribution vectors. The outputs of the two MLPs are summed. The summed vector, which is shown in Figure 2 as a colored vector, is the learned contribution of each $H \times 1$ vector. The vector is called the attention and is used for weighting each $H \times 1$ vector. The weighted $H \times C$ feature map is passed to the next

module. By weighting the 1D-CNN features using the $1 \times C$ attention vectors, the features could be better fused.

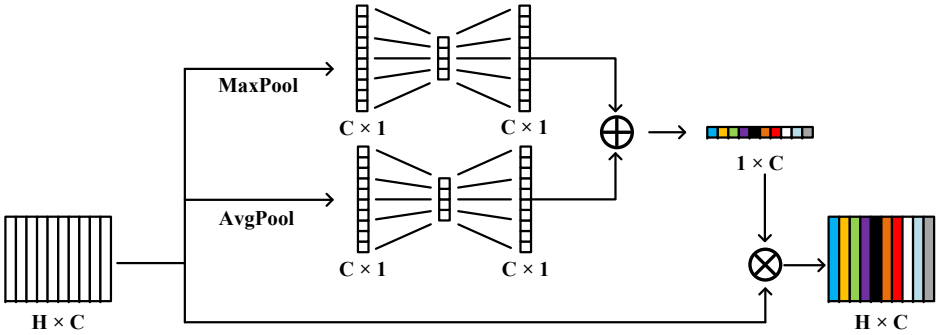



**Figure 2.** The attention module in the proposed approach. The $H \times C$ feature map is firstly reduced to one dimension via two pooling. The two pooling 1D vectors are then inputted to two separate fully connected NNs, respectively. The NNs are used for contribution learning. The outputs of the two NNs are summed to obtained the attention vectors, and are used for feature map weighting.

#### 2.2.3. Prediction Module

In this module, the learned feature maps are projected to a scalar, which is the predicted load at a certain time stamp. Suppose that the dimension of the input feature map is $H \times C$. The input is first pooled to $1 \times C$ by average pooling. The $1 \times C$ features are then weighted and summed to $1 \times 1$ via a fully connected (FC) network, which would be flattened to a scalar. Additionally, the weights in the FC layer are learnable.

### *2.3. Model Implementation Details*

We discuss the details for model implementation in this subsection, including the dimensions of input and output, hyper-parameters of each modules and layers. Suppose that there are $K$ measurement points in the system. The discrete sampled accelerations from a time window, e.g., from $t$ to $t + N\Delta t$, are inputted to the network. We regard the input as $N$ different 1D vectors of $K \times 1$ dimensions. In practice, we set $N = 256$, and $K = 6$. For the two 1D convolution layers in the network, the size of kernel $M$, the number of kernels (or the filter number) $N_{fn}$ , the dilation rate $d$ are given in Table 1. The stride $s$ defines the step size of the kernel when traversing the inputs. Considering $K$ is small in our system, we choose to use a small kernel size $M$. As for the activation function used in the two 1D convolution layers, although we choose the GELU function, the choice for activation function is not limited. The ReLU or Sigmoid function can also be utilized. It is noteworthy that the impact on algorithm performances and metrics with the three activation functions is limited.

**Table 1.** Details for the 1D convolutions in the two convolution modules. In the table, "Conv1D #1" means the first 1d convolution module, and "Conv1D #2" the second.

| Module | Kernel Size $M$ | Filter Number $N_{fn}$ | Dilation Rate $d$ | Stride $s$ |
|---|---|---|---|---|
| Conv1D #1 | 2 | 32 | 2 | 1 |
| Conv1D #2 | 2 | 64 | 2 | 1 |

As for the attention module, recall that two branches are designed, in which the MLPs are used for learning the attention, respectively. The MLP is a fully connected NN with one hidden layer. The number of nodes, which are also known as the neurons, in the hidden layer are set to be identical for the two branches. Table 2 shows the number of the neurons in the input, hidden, and output layer of the MLPs in both attention modules in the network. The ReLU function is used for activation function in the MLPs.

**Table 2.** Details on the amount of neurons for the MLPs in the two-attention module.

| Module | Input Layer | Hidden Layer | Output Layer | Activation |
|---|---|---|---|---|
| ATT #1 | 512 | 1024 | 512 | ReLU |
| ATT #2 | 1024 | 2048 | 1024 | ReLU |

With the above-mentioned input and hyper-parameters of the network layers, the dimensions of the feature maps of all layers can be determined, which is shown in Table 3. The output of the network would be a scalar, which is the predicted load at time $t$. It can be seen from Table 3 that two-convolution and two-attention modules are used in the proposed network. It is noteworthy that more copies of the two modules could be added subsequently between the input layer and prediction module, and the total amount of parameters, computational resources for model training and inference are thus increased.

**Table 3.** The structure of the proposed network. The dimensions of the input and the output feature map of each layer are listed.

| Module | Layer | Input Shape | Output Shape |
|---|---|---|---|
| INPUT | Input | (6.256) | – |
| CONV1D #1 | Conv1D | (6.256) | (4.32) |
| | BN | (4.32) | (4.32) |
| | GELU | (4.32) | (4.32) |
| ATT #1 | Attention | (4.32) | (4.32) |
| CONV1D #2 | Conv1D | (4.32) | (2.64) |
| | BN | (2.64) | (2.64) |
| | GELU | (2.64) | (2.64) |
| ATT #2 | Attention | (2.64) | (2.64) |
| PREDICTION | AvgPooling | (2.64) | (1.64) |
| | FC | (1.64) | (1.1) |
| | Flatten | (1.1) | 1 |
| OUTPUT | Output | – | 1 |

It is noteworthy that the dimension of input and output is not strictly limited. By changing the channels, i.e., $K$, of inputs, the proposed method can also be utilized for load identification problems with less or more than six channels. As well, the dimensions of the output can also be altered, which can be used in MIMO identification problems. The flexibility of the proposed method in terms of input and output dimensions is verified in the experimental results.

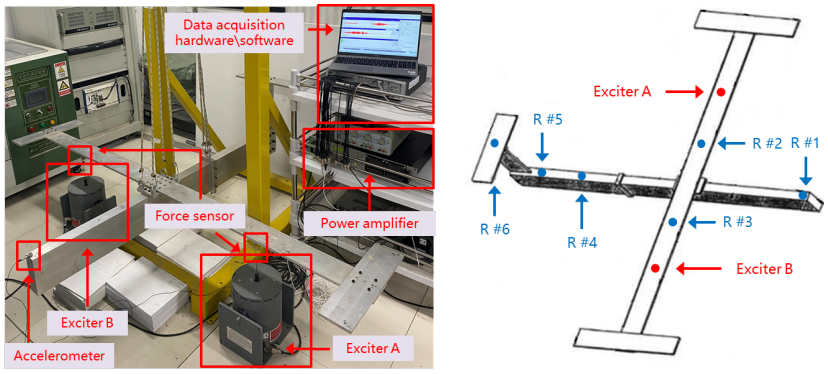
## 3. Experimental Results

In this section, we describe the details of the experimental results for the proposed approach. The equipment and procedure for data collection are given in Section 3.1. Details of model training are described in Section 3.2. The quantitative evaluations and comparisons are presented in Section 3.3. An ablation study of the proposed approach is also presented in Section 3.4.

### *3.1. Data Collection*

To obtain the data for model training and evaluation, in this paper, a self-built GARTEUR aircraft model is utilized for data collection. The GARTEUR aircraft model is a typical and widely used [34–37] standard aircraft model with high compliance, low frequency, and dense frequency characteristics. It is designed to evaluate the accuracy and validity for modal testing methods. The GARTEUR aircraft model consists of six aluminum beams with rectangular section, which are the fuselage, wings, vertical tail, horizontal tail,

and wing end counterweight plates. The fuselage is 1.5 m long and the wingspan is 2.0 m. The GARTEUR model is made of 2024-T3 aluminum alloy, with the elastic modulus of 73 GPa, the density of 2780 kg/m$^3$ and the Poisson's ratio of 0.33. The total weight of the model is 43.34 kg. Each component is connected by screw connection, and an adapter plate is added to connect the fuselage and the wing. Figure 3 shows the self-built GARTEUR aircraft model when collecting experimental data. During all experiments, deformation of the GARTEUR aircraft model is within the range of elastic deformation. The system would be a elastic system satisfying Hooke's law.

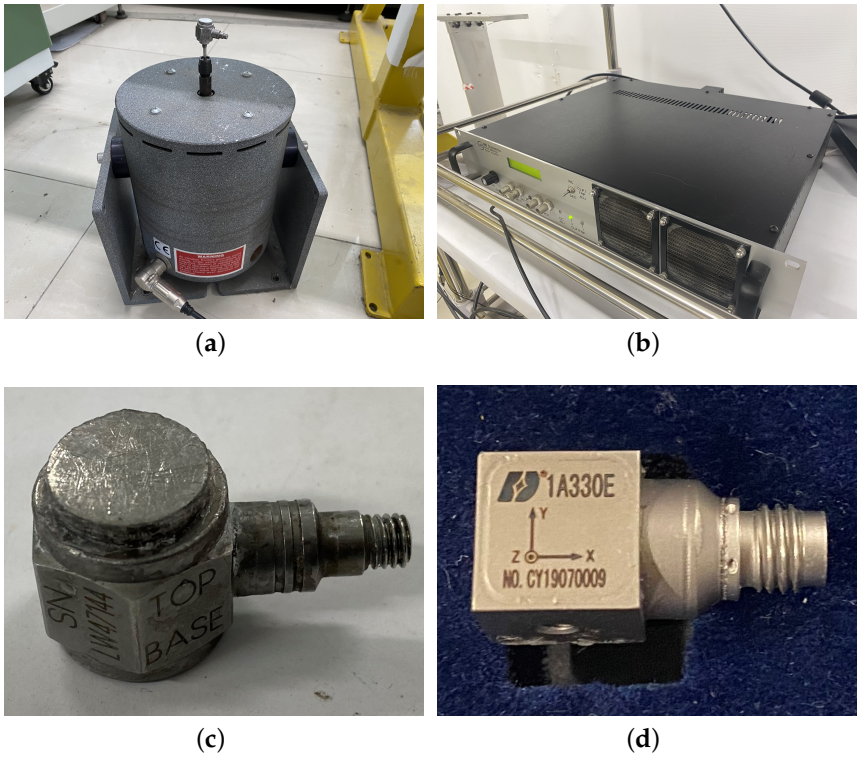



(**a**) (**b**)

**Figure 3.** The self-built GARTEUR aircraft model and equipment for data collection. In the system, two exciters with two force sensors, six response accelerometers are used. The system is synchronized by the data acquisition hardware. The positions of two exciter and six acceleration measurement points are also given. (**a**) The self-built GARTEUR aircraft model used for data collection. (**b**) The positions of exciters and measurements points (R in blue).

To capture the dynamic load and the corresponding response signals, a testing system is built. The entire system is shown in Figure 3a, which includes the vibration exciters, the power amplifiers, the accelerometers, the force sensors, the data acquisition hardware and software. The data acquisition hardware is LMS SCADAS III from Siemens, which is also utilized for signal synchronization.

For the dynamic load signals, the amplified electrical signals are sent to the exciters to excite the GARTEUR aircraft model and vibrate through the exciter rod. In the system, two exciters are used. Exciter A (Figure 3a) provides the random dynamic loads. Random dynamic loads from exciter B (Figure 3a) are considered as the noise load signals for the system. Two high-precision force sensors are placed at the location of the two exciter rods for measuring the loads, respectively. For the corresponding response signals, six high-precision accelerometers are attached to the GARTEUR aircraft model at different locations. Figure 3b shows the locations of sensors for the dynamic loads, the noise loads, and the responses. The accelerometers are placed at the blue points, which are the nose (R #1), the left wing (R #2), the right wing (R #3), the middle of the fuselage (R #4), the rear of the fuselage ((R #5)), and the horizontal tail (R #6). Acceleration measurements from the sensors are captured and stored via data acquisition hardware. The dynamic load signals, the noise load signals, and the six response signals are temporally synchronized via the data acquisition hardware. The exciters, the power amplifier, the force sensors, and the accelerometers utilized in the system are shown in Figure 4.




(**a**) (**b**)

(**c**) (**d**)

**Figure 4.** Equipment used in the data collection system. The dynamic loads are from the vibration exciter and amplified using the power amplifier. The dynamic loads are measured by the force sensor. Responses from each measurement points are measured by the accelerometers. (**a**) The vibration exciter. (**b**) The power amplifier. (**c**) The force sensor. (**d**) The accelerometer.

Throughout the experiment, the random signals of both exciters are white noise. The signal generation module of "LMS Test.lab" from Siemens [38] is utilized to generate the random excitation. The maximum excitation voltage is 0.4 V.

To verify the performance of the proposed method under different noise strengths, the amplitude of load from the exciter B is set to be 0%, 5%, 10%, 20%, 40%, 60%, 80%, 100% of that from the exciter A, respectively. For each configuration of noise strength, 16 s of vibration signals is captured. Since the sampling frequency is 2048 Hz, there would be 32,768 data points for each device (exciter or accelerometer) in one noise strength configuration. Figure 5 shows an example for the dynamic loads, the noise loads, and the responses, with the noise strength of 40%.

In order to determine the sampling frequency, we conducted a modal experiment. By utilizing PolyMAX [39], the modal data between 6.062 Hz (mode 1) to 140.024 Hz (mode 14) of the GARTEUR aircraft model are obtained, which is consistent with that of [40]. Figure 6 shows the modal assurance criterion plot. We choose to set the sampling frequency to 2048 Hz, which is more than 10 times that of the mode 14.

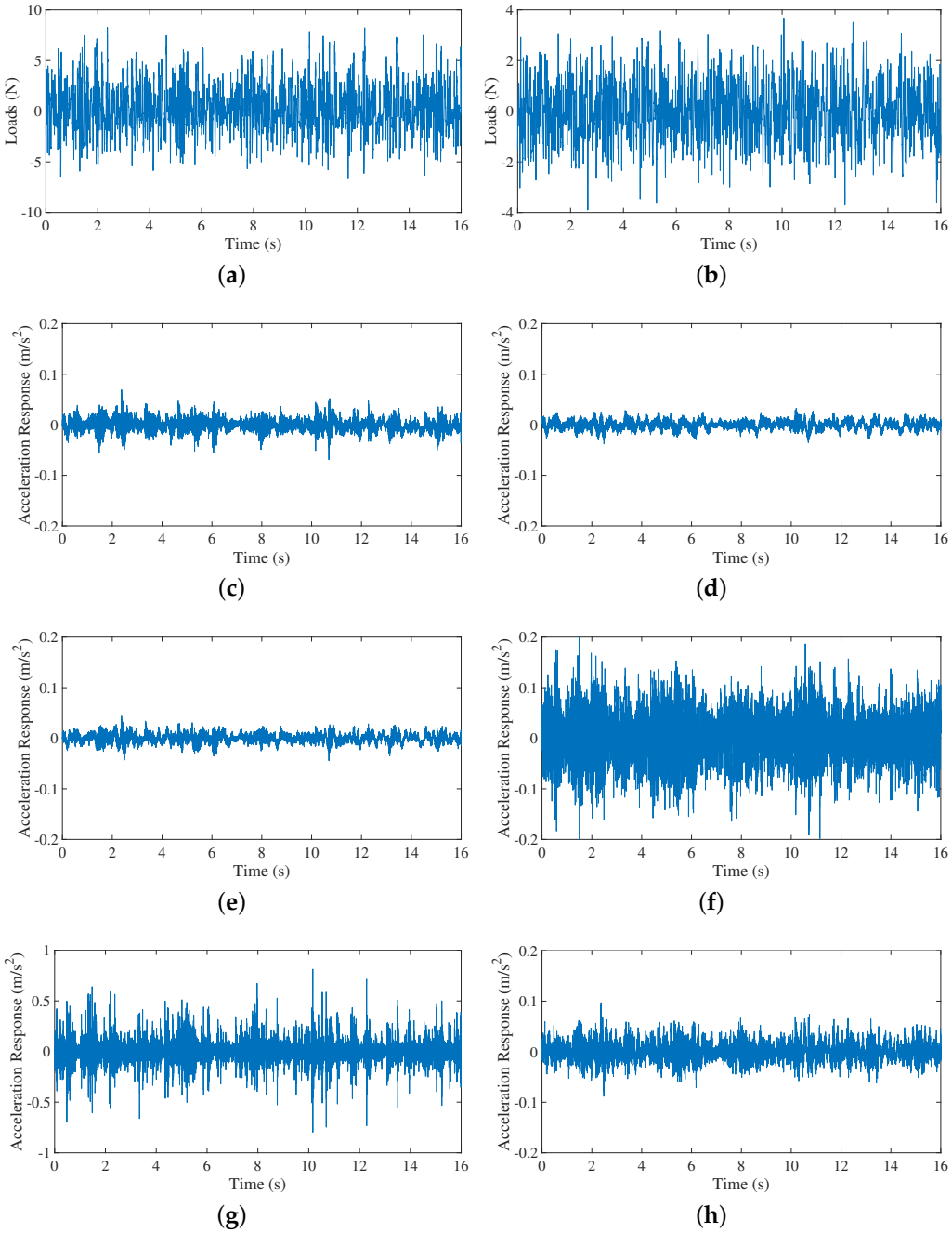



**Figure 5.** An example of the dynamic loads from exciter A and B, and the corresponding responses from 6 measurement points. Loads from exciter B are considered as the noises, in this example, the noise strength, which is the ratio between the amplitudes of loads from exciter A and those from exciter B, is 40%. (**a**) Dynamic loads from exciter A. (**b**) Dynamic loads from exciter B. (**c**) Response from measurements R #1. (**d**) Response from measurements R #2. (**e**) Response from measurements R #3. (**f**) Response from measurements R #4. (**g**) Response from measurements R #5. (**h**) Response from measurements R #6.

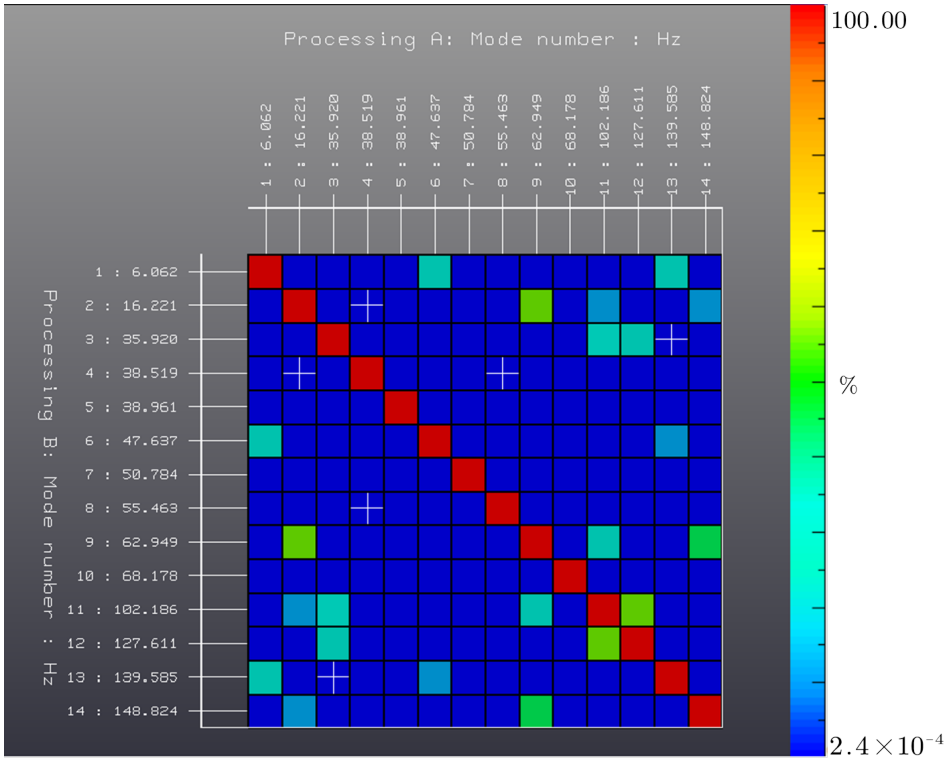



**Figure 6.** The modal assurance criterion plot of 14 modes of the GARTEUR model in our modal experiment. Modal data between 6.062 Hz (mode 1) to 140.024 Hz (mode 14) of the GARTEUR aircraft model are obtained. The sampling frequency is set to 2048 Hz, which is more than 10 times that of mode 14.

### *3.2. Model Training*

For network training, the mean square error ($MSE$) between the ground truth loads $\mathbf{A}$ and identified loads $\mathbf{B}$ is used for the loss function, which is defined as

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{A}_i - \mathbf{B}_i)^2, \tag{7}$$

where $N$ is the amount of data used for training.

We split each 16-second vibration data point into training and testing sets with a proportion of 70% for training and that of 30% for testing. The training and testing is done on a platform of a CPU (Intel i5-10500), and a single GPU (Nvidia RTX 3060 Ti with 8 GB memory). For each experiment, the model is trained with 10 epochs, and the total time for model training on the above mentioned platform is about 4 min. The batch size for training is 1024. The Adam optimization is used for network training with a learning rate of $1.0 \times 10^{-4}$.

### *3.3. Experimental Results*

To quantitatively evaluate the accuracy and effectiveness of proposed approach, several different metrics are utilized for evaluation. The square root of $MSE$, i.e., the $RMSE$, the correlation coefficient ($R$), the relative error ($RE$), the mean absolute error ($MAE$), and the mean absolute percentage error ($MAPE$) are used. Suppose that $\mathbf{A}$ denotes the ground truth of the dynamic load, and $\mathbf{B}$ is the identified dynamic load. $N$ denotes the length of vectors $\mathbf{A}$ and $\mathbf{B}$. The definitions of the above metrics are given below.

$$R = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{\mathbf{A}_i - \boldsymbol{\mu}_A}{\sigma_A}\right)\left(\frac{\mathbf{B}_i - \boldsymbol{\mu}_B}{\sigma_B}\right), \tag{8}$$

where $\boldsymbol{\mu}_A$ and $\sigma_A$ are the mean and standard deviation for the ground truth of the dynamic load, respectively, $\boldsymbol{\mu}_B$ and $\sigma_B$ those of the identified dynamic load.

$$RE = \frac{\|\mathbf{A} - \mathbf{B}\|_2}{\|\mathbf{A}\|_2}, \tag{9}$$

where $\|\cdot\|$ denotes the $L_2$ norm of a vector.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \mid \mathbf{A}_i - \mathbf{B}_i \mid, \tag{10}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \mid \frac{\mathbf{A}_i - \mathbf{B}_i}{\mathbf{A}_i} \mid . \tag{11}$$

where $|\cdot|$ denotes the absolute value.

The experiments are designed as follows. Firstly, the proposed approaches is quantitatively analyzed for inputs with different strength of load noise. Secondly, inputs with different time windows are quantitatively analyzed. Finally, we compare the propose method with conventional regularization-based methods and NN-based approach with quantitative evaluations.

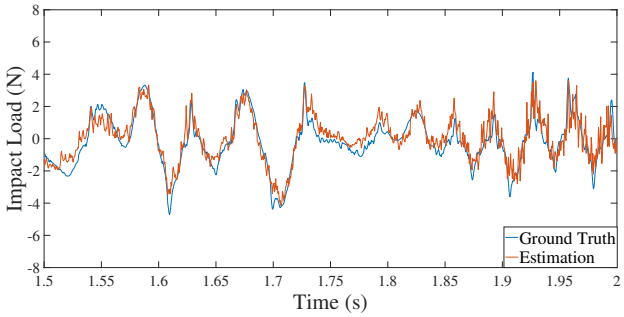
Experiment 1. Different strengths of load noises.

In this experiment, we verify the performance of the proposed approach for different strengths of noises. Recall that the dynamic loads from exciter B would be regarded as noises. The amplitude of loads from exciter B would thus be the strength of the noise. In this experiment, the noise strengths range varies from 0% to 100%. Table 4 presents the quantitative evaluations on testing data (30% of 16-second data) on the above mentioned metrics. In the table, the ↑ sign next to the metric indicates that the larger the metric, the better the performance of the method, while the ↓ sign means the opposite. From Table 4, it can be seen that an increase of noise strength would result in a decrease in algorithm performance. However, the correlation coefficient ($R$) is relatively high in all rows in Table 4, which means the identified loads are consistent with those of the ground truth.

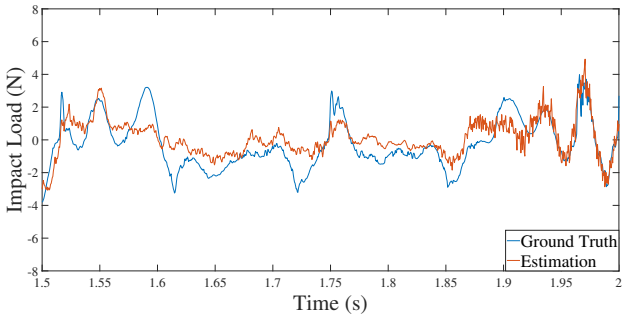
**Table 4.** Quantitative analysis on different strengths of load noises.

| Noise Strength | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|
| 0% | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |
| 5% | 0.7037 | 0.3763 | 0.9272 | 0.5544 | 3.2146 |
| 10% | 0.7637 | 0.3876 | 0.9223 | 0.5935 | 2.1172 |
| 20% | 0.8950 | 0.4235 | 0.9074 | 0.7028 | 1.9737 |
| 40% | 1.0261 | 0.5144 | 0.8599 | 0.8027 | 2.2846 |
| 60% | 1.2355 | 0.5974 | 0.8029 | 0.9550 | 3.0470 |
| 80% | 1.3087 | 0.5982 | 0.8022 | 1.0155 | 2.0965 |
| 100% | 1.2517 | 0.6479 | 0.7628 | 0.9774 | 3.3292 |

Figure 7 gives the comparisons between the identified dynamic loads and the ground truth values under 0% and 40% noise strengths for 0.5 s (from 1.5 s to 2 s). Since the noise strengths are different, the ground truth values in the two sub-figures are different. It can be seen from Table 4 and Figure 7 that the proposed approach can accurately identify the dynamic loads when the noise strengths are low. When the noises are strong, the proposed approach can estimate a reasonable identification results, e.g., $R = 0.7628$ when 100% noise is added. The accuracy of the results in Table 4 would show the potential application of the proposed approach in real aircraft fault diagnoses.




(**a**)




(**b**)

**Figure 7.** Comparisons between the identified random loads and the ground truths under different noise strengths. (**a**) Comparisons with respect to the ground truths (blue) at noise strength of 0% for the proposed method (red). (**b**) Comparisons with respect to the ground truths (blue) at noise strength of 40% for the proposed method (red).

Experiment 2. Different time windows of input signals.

Recall that the load signals and measurements signals are all captured at 2048 Hz. For the proposed method, measurements from a $N = 256$ time window are inputted to the network, which would correspond to 0.125 seconds. In this experiment, we verify the performances of the proposed approach under different input time windows, ranging from $N = 64$ to $N = 384$. Table 5 gives the quantitative results on different metrics under the noise strengths of 0%, 40%, and 100%, representing three cases of no noise, medium strength of noise, and high strength of noise.

**Table 5.** Quantitative analysis on different time windows as inputs.

| Noise Strength | *N* | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|---|
| 0% | 64 | 1.2266 | 0.6352 | 0.7739 | 0.9422 | 2.6046 |
| | 128 | 0.9709 | 0.5028 | 0.8647 | 0.7573 | 2.9802 |
| | 256 | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |
| | 384 | 0.7599 | 0.3918 | 0.9202 | 0.5954 | 2.4893 |
| 40% | 64 | 1.3207 | 0.6736 | 0.7412 | 1.0090 | 2.6030 |
| | 128 | 1.2408 | 0.6334 | 0.7745 | 0.9623 | 2.8295 |
| | 256 | 1.0261 | 0.5144 | 0.8599 | 0.8027 | 2.2846 |
| | 384 | 1.0295 | 0.5160 | 0.8571 | 0.8013 | 2.8821 |
| 100% | 64 | 1.4945 | 0.7570 | 0.6557 | 1.1597 | 3.6410 |
| | 128 | 1.4390 | 0.7335 | 0.6859 | 1.1182 | 3.4713 |
| | 256 | 1.2517 | 0.6479 | 0.7628 | 0.9774 | 3.3292 |
| | 384 | 1.2838 | 0.6698 | 0.7503 | 1.0014 | 3.4256 |

It can be seen from Table 5 that, the size of time window is important to the performances of the algorithm. A larger input time window, which contains more temporal information for the responses, would improve the performances for load identification.

However, when the time window is larger than 256, the improvements would be limited. Therefore, in the other experiments, we choose $N = 256$ as the network input. The results of the experiment would show that, for the proposed algorithm and the experimental setup in the paper, information from an input time window of 0.125 second would be enough for estimating the load for one time stamp.

Experiment 3. Quantitative comparisons of different approaches.

In this experiment, we quantitatively compare the proposed approach with other methods. We choose conventional load identification method and a data driven method for the comparisons.

For the conventional methods, we adopt the TSVD based and Tikhonov regularization approaches. The two methods are both regularization based methods. The TSVD method would truncate small singular values in the transfer function matrix, avoiding disturbance terms generated from small singular values. The Tikhonov regularization method regards the load identification problem as a minimization problem with a residual term and a weighted $L_2$-norm regularization term, which can be explicitly solved. The regularization parameters for both the two approaches are obtained from the GCV method [41]. The two methods are implemented by Matlab.

For the data driven method, the TDNN [26] and a wavelet-LSTM [28] based method are adopted. The TDNN would take the response signals of six measurement points in a time window $[t, t + N\Delta t]$ as the input. The data is processed by a MLP with one hidden layer. The output would be the dynamic load at time $t$. The TDNN is applied for dynamic load identification for aircraft vertical tail model [26]. In [26], a 3-layer NN is designed with the input of time window of $[t, t + 50\Delta t]$, the hidden layer includes 64 neurons. For a reasonable comparison, $N$ is set to be 256, which is the same as that of the proposed method. The learning rate ($1.0 \times 10^{-4}$) and the number of epochs (10) are also the same with those of the proposed method. The TDNN is implemented by Pytorch. For the wavelet-LSTM [28], the input of the network and the output is the eighth-order Meyer wavelet of the signals. The network is with 2 LSTM layer with 128 neurons and 1 FC layer.

Table 6 gives the quantitative comparisons among the four methods. It can be seen that, our methods would outperform in most of the metrics. A higher correlation coefficient ($R$) from the propose method shows that a more consistent identified load with the ground truths. From the comparisons with the TDNN method, it can be seen that the designed network architecture with small but deep convolutional kernel and attention modules in the paper would make more use and fuse better the information of the multiple channel inputs. The proposed method can also outperform LSTM based method. In the meantime, our method avoids the explicit estimation of transferring function and manually tuning of the regularization parameters in the two traditional methods, but resulting in a better identification result. Figure 8 gives the comparisons for different methods with respect to the ground truth values under noise strengths of 100% for 0.5 s (from 1.5 s to 2 s). In all sub-figures of Figure 8, the blue curves are the ground truth values for the random loads, where the identified loads using the TSVD method, those of the Tikhonov method, those of the TDNN method and those of the proposed methods, are shown in magenta, green, yellow, and red, respectively. Figure 8 shows visually that the identified load from the propose method is more consistent with the ground truths, which would be useful for identification problems in real situations.

**Table 6.** Quantitative comparisons on different methods.

| Noise Strength | Method | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|---|
| 0% | TSVD | 1.9392 | 1.0088 | 0.6385 | 1.5373 | 5.2440 |
| | Tikhonov | 1.6975 | 0.8832 | 0.6436 | 1.3345 | 4.1897 |
| | TDNN | 1.5566 | 0.8007 | 0.6547 | 1.2264 | 2.5463 |
| | wavelet-LSTM | 0.8643 | 0.4476 | 0.8967 | 0.6867 | 2.4251 |
| | ours | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |
| 40% | TSVD | 2.4233 | 1.2330 | 0.4788 | 1.9186 | 7.4397 |
| | Tikhonov | 2.4487 | 1.2458 | 0.4695 | 1.9434 | 7.0369 |
| | TDNN | 1.7796 | 0.8922 | 0.4880 | 1.3969 | 2.1664 |
| | wavelet-LSTM | 1.1061 | 0.5647 | 0.8301 | 0.8790 | 2.2342 |
| | ours | 1.0261 | 0.5144 | 0.8599 | 0.8027 | 2.2846 |
| 100% | TSVD | 2.7951 | 1.4172 | 0.3582 | 2.2450 | 9.1424 |
| | Tikhonov | 2.6019 | 1.3230 | 0.3585 | 2.0703 | 8.4509 |
| | TDNN | 1.8333 | 0.9495 | 0.3233 | 1.4649 | 2.5837 |
| | wavelet-LSTM | 1.4504 | 0.7393 | 0.7211 | 1.1689 | 4.0700 |
| | ours | 1.2517 | 0.6479 | 0.7628 | 0.9774 | 3.3292 |

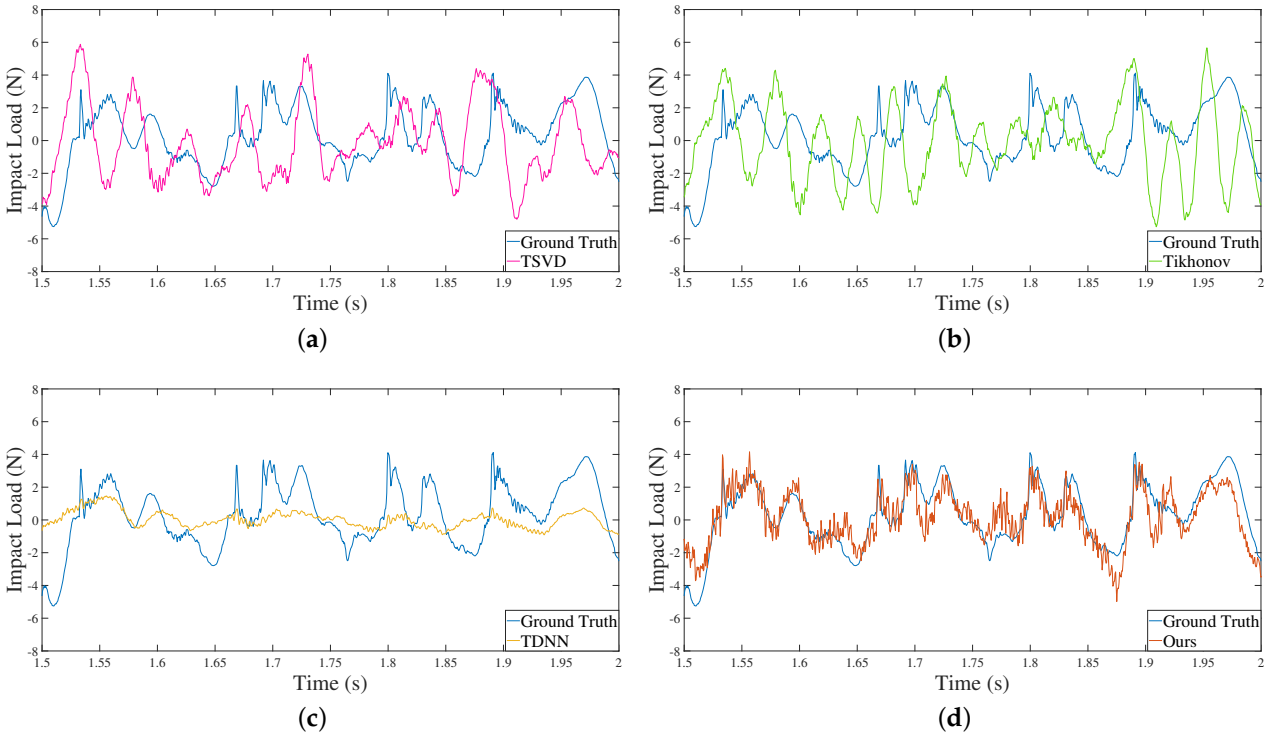



**Figure 8.** The identified dynamic loads and the ground truth values for different input time windows. Four different time windows are utilized in this experiment. The noise strength is 100% in this example. (**a**) Comparisons of the estimated loads using TSVD method (magenta) with respect to the ground truths (blue). (**b**) Comparisons of the estimated loads using Tikhonov method (green) with respect to the ground truths (blue). (**c**) Comparisons of the estimated loads using TDNN method (yellow) with respect to the ground truths (blue). (**d**) Comparisons of the estimated loads using the proposed method (red) with respect to the ground truths (blue).

Experiment 4. Further MISO and MIMO identification performance verification.

In this experiment, we plan to further show the flexibility of the proposed method in terms of dimensions of inputs and outputs. First, by changing the dimensions of input, the proposed method can be used for different MISO configurations. To verify this, we use the 0% noise data for experiment. Table 7 shows the quantitative evaluation results.

**Table 7.** Quantitative comparisons on different numbers of input channels. By changing the dimension of the inputs of the network, the proposed method can fit different MISO configuration.

| Noise Strength | Channel Number | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|---|
| 0% | 1 | 1.5033 | 0.7733 | 0.6450 | 1.1685 | 3.7730 |
| | 2 | 1.5052 | 0.7743 | 0.6447 | 1.1684 | 3.8007 |
| | 3 | 1.2481 | 0.6420 | 0.7796 | 0.9814 | 3.6928 |
| | 4 | 1.1995 | 0.6171 | 0.8001 | 0.9293 | 3.4354 |
| | 5 | 0.7719 | 0.3971 | 0.9187 | 0.6116 | 2.3188 |
| | 6 | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |

From Table 7, it can be seen that the increase of input channels can improve the performances, since more information is provide. Moreover, the designed convolutional modules and attention modules in the proposed network can learns the additional information from extra input channels, resulting in an increase of quantitative metrics.

By comparing the first row of Table 7 with the performance of conventional methods, i.e., the first and second rows of Table 6, it can be seen that the proposed network with one-channel input would perform similar with the conventional methods. However, the explicitly estimation of transfer function matrix can be avoided in our method.

Second, by changing the dimensions of the network output, the proposed method can be used for MIMO load identification. To verify the flexibility, the six accelerometers' signals are used for multiple inputs and the load of both Exciter A and B are estimated. Three sets of data are utilized for evaluation, in which the ratios between the amplitude of Exciter A and that of B are 60%, 80%, and 100%, respectively. Table 8 presents the quantitative analysis results in this experiment. By comparing results in Table 8 with those in Table 4, a performance drop is noticed in estimation of load from Exciter A.

The results in this experiment showed that the proposed method would have the flexibility to fit different configuration of SISO, MISO and MIMO load identification problems. Once again, the explicit modeling of the transferring function from the input to the output can be avoided, while a reasonable identification results can be obtained, which would be meaningful in real-situation applications.

**Table 8.** Quantitative analysis on the MIMO identification performances. By changing the output dimensions in the network, the load of Exciter A and B can be simultaneously estimated.

| Amplitude Ratio | Exciter | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|---|
| 60% | A | 1.2112 | 0.5857 | 0.8122 | 0.9411 | 2.9255 |
| | B | 1.2135 | 0.7742 | 0.6402 | 0.9595 | 3.6395 |
| 80% | A | 1.3496 | 0.6168 | 0.7873 | 1.0493 | 2.1058 |
| | B | 1.4671 | 0.7292 | 0.6844 | 1.1695 | 6.3053 |
| 100% | A | 1.3031 | 0.6745 | 0.7382 | 1.0134 | 3.1291 |
| | B | 1.7538 | 0.7199 | 0.6953 | 1.9758 | 2.8092 |

Experiment 5. Experiments on additional sensor measurement noise.

In addition to the interference from another load, another potential factor that might affect the performance of the algorithm would be the measurement noise from the accelerometers. It should be noted that the data used in all experiments in the paper are from real sensors, therefore, there are already measurement noises in the accelerometers output. The goal of this experiment is to further analyze the algorithm performance under sensor measurement noise.

The data when only Exciter A is working are used for evaluation to avoid other interferences. A time series Gaussian noise with the mean value of 0 to is added to each of the accelerometer measurements. The standard deviation is set to be $p$% of the amplitude of the accelerometer measurement. In the experiment, $p$% varies from 0% to 10%. Table 9

show the performances of the proposed method on different additional sensor noises. The quantitative results in Table 9 shows that the performances of the algorithm would decrease with the increase of the additional sensor noise.

**Table 9.** Quantitative analysis on different strengths of additional sensor noises.

| Sensor Noise Strength | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|
| 0% | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |
| 2% | 0.9506 | 0.4890 | 0.8742 | 0.7434 | 3.0471 |
| 4% | 1.1763 | 0.6051 | 0.8000 | 0.9099 | 3.4878 |
| 6% | 1.3305 | 0.6844 | 0.7356 | 1.0447 | 3.4711 |
| 8% | 1.5692 | 0.8072 | 0.6153 | 1.2275 | 3.6680 |
| 10% | 1.5608 | 0.8029 | 0.6070 | 1.2234 | 4.1296 |

To further the evaluate the performances of the proposed method under sensor noise, the above metrics are compared with the conventional methods, namely, the TSVD method and Tikhonov regularization method. Table 10 shows the quantitative comparison. It can be seen that the drop rate in metric $R$ is larger for our method than for the conventional methods. The proposed method could outperform at a low noise strength. It should be noted that the main motivation of the proposed method is not to suppress the sensor noise. A further design of the network and fine-tuning of the parameters would be useful for a better performance under high sensor noise.

**Table 10.** Quantitative comparisons on different strengths of additional sensor noises.

| Sensor Noise Strength | Method | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|---|
| 0% | TSVD | 1.5094 | 0.7853 | 0.6192 | 1.1615 | 2.9040 |
| | Tikhonov | 1.5448 | 0.8037 | 0.6223 | 1.2041 | 3.4852 |
| | ours | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |
| 6% | TSVD | 1.5730 | 0.8183 | 0.5749 | 1.2047 | 3.0625 |
| | Tikhonov | 1.5920 | 0.8282 | 0.6050 | 1.2413 | 3.6896 |
| | ours | 1.3305 | 0.6844 | 0.7356 | 1.0447 | 3.4711 |
| 10% | TSVD | 1.5960 | 0.8303 | 0.5575 | 1.2265 | 2.9476 |
| | Tikhonov | 1.7099 | 0.8896 | 0.5658 | 1.3276 | 4.1587 |
| | ours | 1.5608 | 0.8029 | 0.6070 | 1.2234 | 4.1296 |

Experiment 6. Sensibility experiment on different sensor positions.

As the responses of the load would differ at different positions on the aircraft structure, the performance of the proposed method may be affected by the positions of the sensors. In this experiment, our goal is to test the potential sensibility of the algorithm in different sensor positions. To avoid other potential interferences, only one exciter is used in the experiment. Figure 9 shows the sensor positions in symmetrical and asymmetrical distributions in the experiment.

Table 11 shows the comparisons of the algorithm performance with inputs from different positions of accelerometers. It can be seen that the performance is different. The case of symmetrical distribution would outperform. From the comparisons, it can be seen that the proposed algorithm is sensitive to the sensor position. Since the complexity of the transfer functions for different positions on the aircraft model is different, the accuracy of the implicit estimation from the neural networks of such "transfer functions" would be different given the same amount of input signals. The differences would make the sensibility of the proposed method.

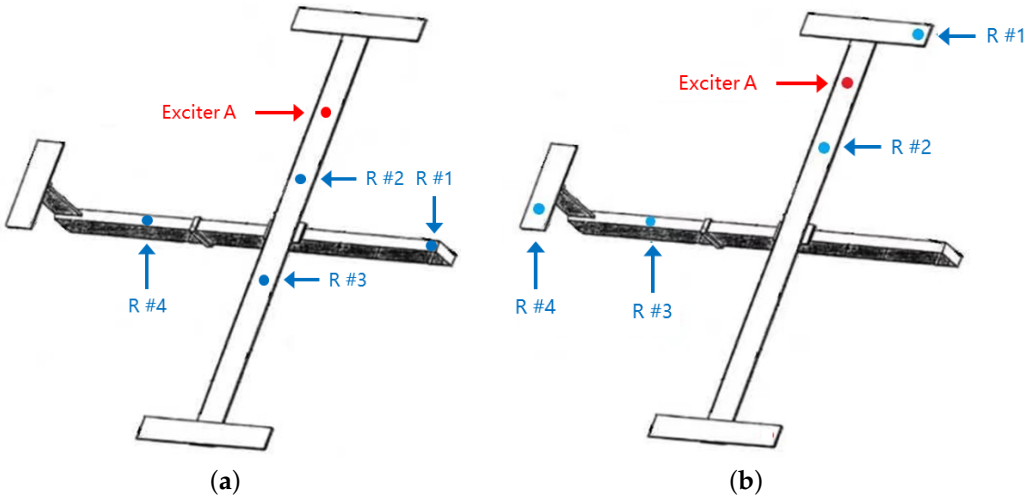



**Figure 9.** Symmetrically and asymmetrically distributed accelerometers in the experiment. In total, 4 accelerometers are used. To avoid other interferences, only 1 exciter is used in exepriment. (**a**) Symmetrically distributed 4 accelerometers and 1 exciter on the aircraft model. (**b**) Asymmetrically distributed 4 accelerometers and one exciter on the aircraft model.

**Table 11.** Quantitative comparisons on different distribution of sensors.

| Accelerometers Distribution | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|
| Symmetrical | 1.1995 | 0.6171 | 0.8001 | 0.9293 | 3.4354 |
| Asymmetrical | 2.0609 | 0.7090 | 0.7079 | 1.6535 | 2.7237 |

### *3.4. Ablation Study*

We perform ablations on the proposed approach to determine the quantitative contribution for load identification of the feature weighting mechanism. The ablations are performed as follows. For the case of "no-att" (no-attention), no attention module would be used in the network, while the dimensions of input, convolution modules, prediction modules would remain. The ablation is performed on the data with 0%, 40%, and 100% noise strength, representing low, mid, and high levels of noises. Table 12 shows the quantitative comparisons on the metrics.

**Table 12.** Ablations on different noise strengths.

| Noise Strength | Method | RMSE ↓ | RE ↓ | R ↑ | MAE ↓ | MAPE ↓ |
|---|---|---|---|---|---|---|
| 0% | no-att | 0.7827 | 0.4026 | 0.9161 | 0.6222 | 2.5564 |
| | ours | 0.7517 | 0.3867 | 0.9232 | 0.5939 | 2.4077 |
| 40% | no-att | 1.0660 | 0.5344 | 0.8531 | 0.8384 | 2.2758 |
| | ours | 1.0261 | 0.5144 | 0.8599 | 0.8027 | 2.2846 |
| 100% | no-att | 1.3040 | 0.6750 | 0.7380 | 1.0202 | 2.8646 |
| | ours | 1.2517 | 0.6479 | 0.7628 | 0.9774 | 3.3292 |

From Table 12, it can be seen that, with the designed feature weighting mechanism, the performance on all metrics would be better than those of the "no-attention" situation. Additionally, the improvement of the metrics would be larger when the noise strengths are larger. The learned attention would help to increase the contribution of more important learned high level features. As for the time for training and testing, for the case of "no-attention", the time for training is 190 s, and the testing would cost 10 s. For the proposed approach, the time for training is 260 s, and the testing time is 20 s.

## 4. Conclusions

In this paper, we have presented a novel attention based 1D-CNN model to identify the dynamic loads for aircraft models. As a data-driven approach, the proposed model would avoid the accurate structure modeling. The proposed approach takes advantage of the 1D-CNN to extract temporal features. The learned features are then weighted using learned attentions. The main limitations for the proposed approach is that response data of a certain period of time are needed for model training. In the future, we plan to further analyze the sensibility of sensor positions, while suppressing the potential high sensor measurement noise. A test of the proposed model on real civil aircrafts, while performing a further exploration on the attention mechanism would be other future directions.

**Author Contributions:** Conceptualization, X.Z.; Data curation, Q.L. and B.X.; Formal analysis, X.Z.; Investigation, B.X.; Methodology, W.H. and Z.F.; Project administration, X.L.; Software, W.H., X.Z., and Q.L.; Supervision, Z.F.; Visualization, Q.L.; Writing—original draft, W.H. and X.Z.; Writing—review and editing, Z.F. and X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported in part by the Open Fund of Key Laboratory of Civil Aircraft Airworthiness Technology under Grant SH2020112706, in part by the Doctoral Fund of Southwest University of Science and Technology under Grant 19zx7123, in part by the Fundamental Research Funds for the Central Universities (special project of the Civil Aviation University of China) under Grant 3122019163.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

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

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| GCV | general cross-validation |
| TSVD | truncated singular value decomposition |
| NN | neural network |
| LSTM | long short term memory |
| MLP | multilayer perceptron |
| CNN | conventional neural network |
| TDNN | time delay neural network |
| 1D-CNN | one-dimensional convolution neural network |
| BN | batch normalization |
| GELU | Gaussian error linear unit |
| ReLU | rectified linear unit |
| FC | fully connected |
| MSE | mean square error |
| RMSE | root of mean square error |
| RE | relative error |
| MAE | mean absolute error |
| MAPE | mean absolute percentage error |
| SISO | single input single out |
| MISO | multiple inputs single output |
| MIMO | multiple inputs multiple outputs |

## References

1. Candon, M.; Esposito, M.; Fayek, H.; Levinski, O.; Koschel, S.; Joseph, N.; Carrese, R.; Marzocca, P. Advanced multi-input system identification for next generation aircraft loads monitoring using linear regression, neural networks and deep learning. *Mech. Syst. Signal Process.* **2022**, *171*, 108809. [CrossRef]
2. Morse, L.; Cartabia, L.; Mallardo, V. Reliability-based bottom-up manufacturing cost optimisation for composite aircraft structures. *Struct. Multidiscip. Optim.* **2022**, *65*, 1–27. [CrossRef]

3. Dangut, M.D.; Jennions, I.K.; King, S.; Skaf, Z. Application of deep reinforcement learning for extremely rare failure prediction in aircraft maintenance. *Mech. Syst. Signal Process.* **2022**, *171*, 108873. [CrossRef]
4. Cartocci, N.; Napolitano, M.R.; Costante, G.; Valigi, P.; Fravolini, M.L. Aircraft robust data-driven multiple sensor fault diagnosis based on optimality criteria. *Mech. Syst. Signal Process.* **2022**, *170*, 108668. [CrossRef]
5. Agrapart, Q.; Nyssen, F.; Lavazec, D.; Dufrénoy, P.; Batailly, A. Multi-physics numerical simulation of an experimentally predicted rubbing event in aircraft engines. *J. Sound Vib.* **2019**, *460*, 114869. [CrossRef]
6. Liu, R.; Dobriban, E.; Hou, Z.; Qian, K. Dynamic load identification for mechanical systems: A review. *Arch. Comput. Methods Eng.* **2022**, *29*, 831–863. [CrossRef]
7. Uhl, T. The inverse identification problem and its technical application. *Arch. Appl. Mech.* **2007**, *77*, 325–337. [CrossRef]
8. Jacquelin, E.; Bennani, A.; Hamelin, P. Force reconstruction: Analysis and regularization of a deconvolution problem. *J. Sound Vib.* **2003**, *265*, 81–107. [CrossRef]
9. Wang, L.; Huang, Y.; Xie, Y.; Du, Y. A new regularization method for dynamic load identification. *Sci. Prog.* **2020**, *103*, 0036850420931283. [CrossRef]
10. Jiang, J.; Tang, H.; Mohamed, M.S.; Luo, S.; Chen, J. Augmented tikhonov regularization method for dynamic load identification. *Appl. Sci.* **2020**, *10*, 6348. [CrossRef]
11. Bianchi, D.; Buccini, A.; Donatelli, M.; Serra-Capizzano, S. Iterated fractional Tikhonov regularization. *Inverse Probl.* **2015**, *31*, 055005. [CrossRef]
12. Bao, Y.; Li, H.; Chen, Z.; Zhang, F.; Guo, A. Sparse $l_1$ optimization-based identification approach for the distribution of moving heavy vehicle loads on cable-stayed bridges. *Struct. Control Health Monit.* **2016**, *23*, 144–155. [CrossRef]
13. Pan, C.D.; Yu, L.; Liu, H.L.; Chen, Z.P.; Luo, W.F. Moving force identification based on redundant concatenated dictionary and weighted l1-norm regularization. *Mech. Syst. Signal Process.* **2018**, *98*, 32–49. [CrossRef]
14. Liu, J.; Qiao, B.; Chen, Y.; Zhu, Y.; He, W.; Chen, X. Impact force reconstruction and localization using nonconvex overlapping group sparsity. *Mech. Syst. Signal Process.* **2022**, *162*, 107983. [CrossRef]
15. Li, Q.; Lu, Q. Time domain force identification based on adaptive $L_q$ regularization. *J. Vib. Control* **2018**, *24*, 5610–5626. [CrossRef]
16. Trivailo, P.; Carn, C. The inverse determination of aerodynamic loading from structural response data using neural networks. *Inverse Probl. Sci. Eng.* **2006**, *14*, 379–395. [CrossRef]
17. Chen, G.; Li, T.; Chen, Q.; Ren, S.; Wang, C.; Li, S. Application of deep learning neural network to identify collision load conditions based on permanent plastic deformation of shell structures. *Comput. Mech.* **2019**, *64*, 435–449. [CrossRef]
18. Wang, J.; Zhang, J.; Wang, Y.; Fang, X.; Zhao, Y. Nonlinear identification of one–stage spur gearbox based on pseudo–linear neural network. *Neurocomputing* **2018**, *308*, 75–86. [CrossRef]
19. Zhou, J.; Dong, L.; Guan, W.; Yan, J. Impact load identification of nonlinear structures using deep Recurrent Neural Network. *Mech. Syst. Signal Process.* **2019**, *133*, 106292. [CrossRef]
20. Hossain, M.S.; Ong, Z.C.; Ismail, Z.; Khoo, S.Y. A comparative study of vibrational response based impact force localization and quantification using radial basis function network and multilayer perceptron. *Expert Syst. Appl.* **2017**, *85*, 87–98. [CrossRef]
21. Gao, Y.; Yu, X.; Chen, L.; Huang, D. Impact Load Identification Algorithm of Helicopter Weapon Pylon Based on Time-Domain Response Signal. *Aerospace* **2022**, *9*, 388. [CrossRef]
22. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107398. [CrossRef]
23. Du, C.; Zhang, X.; Zhong, R.; Li, F.; Yu, F.; Rong, Y.; Gong, Y. Unmanned aerial vehicle rotor fault diagnosis based on interval sampling reconstruction of vibration signals and a one-dimensional convolutional neural network deep learning method. *Meas. Sci. Technol.* **2022**, *33*, 065003. [CrossRef]
24. Guo, D.; Zhong, M.; Ji, H.; Liu, Y.; Yang, R. A hybrid feature model and deep learning based fault diagnosis for unmanned aerial vehicle sensors. *Neurocomputing* **2018**, *319*, 155–163. [CrossRef]
25. Dang, V.H.; Vu, T.C.; Nguyen, B.D.; Nguyen, Q.H.; Nguyen, T.D. Structural damage detection framework based on graph convolutional network directly using vibration data. *Structures* **2022**, *38*, 40–51. [CrossRef]
26. Xia, P.; Yang, T.; Xu, J.; Wang, L.; Yang, Z. Reversed time sequence dynamic load identification method using time delay neural network. *Acta Aeronaut. Astronaut. Sin.* **2021**, *42*, 224452.
27. Yang, H.; Jiang, J.; Chen, G.; Zhao, J. Dynamic load identification based on deep convolution neural network. *Mech. Syst. Signal Process.* **2023**, *185*, 109757. [CrossRef]
28. Yang, T.; Yang, Z.; Liang, S.; Kang, Z.; You, J. Feature extraction and indentification of stationary random dynamic load using deep neural network. *Acta Aeronaut. Astronaut. Sin.* **2022**, *43*, 225952.
29. Oord, A.v.d.; Dieleman, S.; Zen, H.; Simonyan, K.; Vinyals, O.; Graves, A.; Kalchbrenner, N.; Senior, A.; Kavukcuoglu, K. Wavenet: A generative model for raw audio. *arXiv* **2016**, arXiv:1609.03499.
30. Yu, F.; Koltun, V. Multi-Scale Context Aggregation by Dilated Convolutions. In Proceedings of the International Conference on Learning Representations (ICLR), San Juan, Puerto Rico, 2–4 May 2016.
31. Ran, L.; Zhang, Y.; Hua, G. CANNET: Context aware nonlocal convolutional networks for semantic image segmentation. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 4669–4673.
32. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv* **2020**, arXiv:cs.LG/1606.08415.

33. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023. [CrossRef]
34. Balmes, E.; Wright, J.R. GARTEUR group on ground vibration testing: Results from the test of a single structure by 12 laboratories in Europe. In Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Sacramento, CA, USA, 14–17 September 1997; American Society of Mechanical Engineers, New York, CA, USA, 1997; Volume 80401, p. V01AT03A004.
35. Chen, J.; Jiang, J.; Wang, K.; Zhang, F. Optimal Placement of Actuators for Active Vibration Control Using EER and Genetic Algorithm. In Proceedings of the 2019 IEEE 10th International Conference on Mechanical and Aerospace Engineering (ICMAE), Brussels, Belgium, 22–25 July 2019; pp. 449–453. [CrossRef]
36. Bao, N.; Wang, C. A Monte Carlo simulation based inverse propagation method for stochastic model updating. *Mech. Syst. Signal Process.* **2015**, *60–61*, 928–944. [CrossRef]
37. Rett, S.; Nabarrete, A.; Arbelo, M.; Góes, L.C.; Guimarães, G. Results of the GVT of the Unmodified GARTEUR SM-AG-19 Testbed in South America. In Proceedings of the 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 18th AIAA/ASME/AHS Adaptive Structures Conference 12th, Orlando, FL, USA, 12–15 April 2010; p. 2808.
38. Siemens-Simcenter. Simcenter Testlab, formerly called LMS Test.Lab. 2022. Available online: https://community.sw.siemens.com/s/article/simcenter-testlab (accessed on 10 December 2022. ).
39. Peeters, B.; Van der Auweraer, H.; Guillaume, P.; Leuridan, J. The PolyMAX frequency-domain method: A new standard for modal parameter estimation? *Shock Vib.* **2004**, *11*, 395–409. [CrossRef]
40. Link, M.; Friswell, M. Working Group 1: Generation of validated structural dynamic models—results of a benchmark study utilising the GARTEUR SM-AG19 test-bed. *Mech. Syst. Signal Process.* **2003**, *17*, 9–20. [CrossRef]
41. Hansen, P.C. Regularization tools version 4.0 for Matlab 7.3. *Numer. Algorithms* **2007**, *46*, 189–194. [CrossRef]