ORCA – Online Research @ Cardiff

Self-validating high-g accelerometers through data-driven methods

Jingjing Wen[a*], Houpu Yao[b], Ze Ji[c*], Bin Wu[a], Feng Xu[a*]

([a.] School of Astronautics, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China;

[b.] JD Finance America Corporation, Mountain View, CA 94043, USA;

[c.] School of Engineering, Cardiff University, Cardiff CF243AA, UK.)

[*.] Correspondence: jiz1@cardiff.ac.uk (Z. Ji); xufeng@nwpu.edu.cn (F. Xu); wjj1990@mail.nwpu.edu.cn (J. Wen).

Abstract: In the aerospace industry, pyroshock testing is an indispensable step in designing space electronics. Yet, damages in high-g accelerometers, the core measuring instruments in pyroshock test systems, could result in various failures of pyroshock tests. To ensure the reliability of pyrochock tests for space electronics, a machine learning system is proposed to perform self-validation for high-g accelerometers. In this work, self-validation refers to the capability of identifying five key parameters, namely the validated shock signal, the validated uncertainty, the measurement status, the raw shock signal and the fault type, synchronously during measuring shock signals in pyroshock tests. To achieve the highest performance, we accomplish these tasks through combining an ensemble learning model and a deep neural network (DNN). The ensemble learning model, which integrates several k-nearest neighbors with different k values, is used to identify the sensors' health conditions from their measurements and diagnose their fault types synchronously if damaged. The DNN, a deep autoencoder-based neural network, is designed to correct corrupted measurements through constructing the mapping between faulty signals and their corresponding reference counterparts. Experimental results show that the proposed machine learning system is capable of not only accurately identifying the health conditions and fault types of the damaged high-g accelerometers from their measurements, but also recovering the corrupted shock signals to a large extent, and, meanwhile, outputting the five self-validation parameters.

Key words: pyroshock test; high-g accelerometer; self-validation sensor; ensemble kNNs; deep stacked autoencoders

1. Introduction

Pyroshock events, including the release of payloads, separation of launcher stages, the deployment of solar panels, the activation of impulse engines etc., will occur repeatedly throughout aerospace missions [1]. Although pyroshock causes less damage on structural parts of space devices, it could lead to various malfunctions in electronic and/or optical components, which would further result in partial or total failures of flight missions [2]. To evaluate the reliability and survivability of space electronics under pyroshock environments, many authority bodies, such as NASA, JEDEC, and ISO, have proposed several ground test specifications, and passing these specifications is required mandatorily during designing and/or manufacturing aerospace electronics [3].

In a pyroshock testing system, high-g accelerometer is one of the pivotal tools widely applied to accurately measure shock signals [4]. However, in practice, high-g accelerometers are likely to be damaged under harsh pyroshock environments [5]. Consequently, the damaged high-g accelerometers will measure erroneous shock data and thereby failing the pyroshock test. A conservative solution for such problems is to discard all the erroneous data and repeat the pyroshock test by replacing the faulty accelerometers [6]. Under many circumstances, however, the acquired shock data is unique, and repeating the pyroshock test can be extremely difficult or costly [7]. A typical example is that when the to-be-tested electronic device and the accelerometer are both damaged at the same time during the high-g shock test [8]. Unfortunately, it is quite costly to repeat this destructive test, and, meanwhile, the faulty accelerometer cannot provide accurate measurements to help analyzing the damage causes of the

defective electronic device. Another representative example is that an uncalibrated accelerometer is used to measure and collect massive shock data, but without awareness of the missing calibration beforehand [9]. Also, it is considerably difficult to repeat these shock tests because of the huge resource waste. Additionally, Yuan et al. [10] also present a practical case that mass shock data telemetered from a rocket flight test contains spurious noises caused by faulty sensors. Similarly, these noise-polluted shock data cannot be used to evaluate the health condition of the rocket, and it is impermissible to repeat the flight test either. Therefore, it would be of great interest for both the end-users and manufacturers to develop an effective health management system that can monitor the health conditions of the accelerometers in use; diagnose the fault types of the damaged sensors; and restore the desired data from the abnormal readings provided by the faulty accelerometers synchronously.

To address the above mentioned challenges, Henry and Clarke from Oxford University proposed a self-validating methodology for sensors in 1993 [11], of which the function framework can be summarized in Fig. 1. It can be seen form Fig. 1 that, compared with traditional sensors, the fault diagnosis module and the output data generation module are compensated in the function framework of self-validating sensors. These two modules should guarantee the self-validating sensors to have the abilities to 1) self-diagnose in real time and output the health conditions and fault types (if the sensors failure is detected) timely; 2) reconstruct the measurements of the faulty sensors within a certain accuracy range and replace the original measurements with the reconstructed results; and 3) estimate the uncertainties of the output measurements.
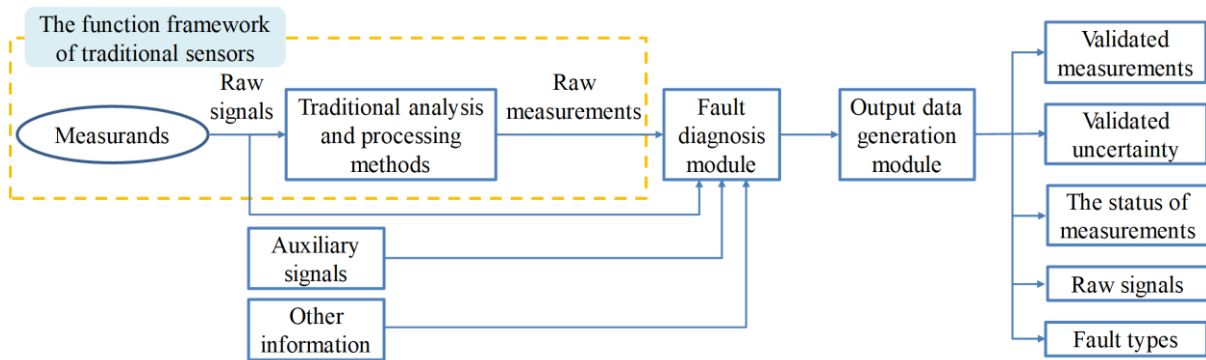

Fig. 1. The function framework of the self-validating sensor and the traditional sensor. This figure is adapted from [12].

Based on the above requirements, five output parameters were defined for self-validating sensors, including raw measurements, validated measurements, validated uncertainty, the status of measurements, and the fault types. Raw measurements are the output signals of the sensors without self-validation. If a sensor is healthy, the raw measurements are the desired signals, and, otherwise, the raw measurements will be faulty signals. Validated measurements refer to the output signals of the sensors after self-validation. If the sensor is healthy, the validated measurements would be the same as the raw measurements using healthy sensor directly, while, if the sensor is faulty, the validated measurements would be the reconstructed signals, and, we expect that the reconstructed signals could be as close to the desired signals as possible. Validated uncertainty is the measure to quantify the effect of a fault on the raw measurements. Obviously, the validated uncertainty will be higher with more severe sensor failures in presence. The status of measurements reflects the working conditions of the sensors and can inform the users if the measurements were carried out under normal or failure statuses. If the sensor is faulty, the self-validation system should output the detailed fault type for the users.

Based on Henry and Clarke's definition, researchers make appropriate adjustments on these five output parameters for specific problems, so as to design suitable self-validation sensors for corresponding application scenarios. Regarding to multifunctional self-validating sensors, failure detection, isolation, and data recovery are of high concern [13-15]. In the field of self-validating multifunctional sensors, Yang et al. [15] develop an efficient approach by upgrading traditional contribution plots method for self-validating multifunctional sensors, which

consist of metal-oxide sensor, temperature sensor, and humidity sensor. This approach has the advantage of isolating multiple faults accurately but needs to calculate the fault direction set and the fault amplitude set in advance. For self-validating wireless sensor networks, the emphasis is placed on the local uncertainty evaluation [16]. For example, Liu et al. [17] adopt a low-rank completion theory-based algorithm to correct measuring errors in the passive wireless SAW resonant sensor system. This approach gets rid of the reliance on system modeling, probability distribution assumption, and redundant sensors, but cannot identify the accurate health conditions of sensors. Besides, this method requires distinct frequency difference, meaning that this approach is more suitable for recovering sensors with low-varying parameters like temperature sensors. Additionally, for self-validating pressure sensors, fault detection, fault diagnosis, and data recovery are the main focuses [18]. In [18], Feng et al. apply an LS-SVM-based predictor to recover a pressure sensor's faulty measurement. This method mainly aims at correcting pressure values at corresponding sampling points but lacks the ability to recover signals with long durations and/or high sampling rates. More recently, soft sensors have attracted greatly increasing attentions from researchers for industrial process monitoring. However, the development of soft sensors presents many challenges using conventional model-driven approaches and have been greatly inspired and facilitated by techniques derived from machine learning or data driven methods. Data-driven self-validation is one of the immediately identified open challenges for future research [19]. On the other hand, key-performance-indicators (KPI)-oriented prognosis and fault diagnosis provide a mechanism to understand the operation states and provide a guide for the subsequent fault maintenance. It will be highly desired for self-validation to be able to distinguish how a local malfunction affects the systems KPIs [20].

As aforementioned, extensive works have been performed on temperature sensors, gas detectors, and pressure sensors, but to the best of our knowledge, self-validation on high-g accelerometers has not been studied much. Previous explorations are limited to simple failure types like shock spectrum distortion and baseline drift [6, 7, 21, 22], and are not applicable to failures of more complicated forms. In these works, faulty signals are identified and corrected on the basis of hard coded rules in signal processing, including the wavelet transform, empirical mode decomposition, Fourier analysis, etc. In addition, differing from the measuring results from the aforementioned sensors, pyroshock signals usually contain strong nonlinearity and uncertainty, reflected in their complicated frequency spectrums, drastically varying amplitudes, and much shorter durations [23]. Meanwhile, these nonlinear and uncertain characteristics also increase the difficulty to self-validate the measurements of high-g accelerometers. Firstly, the self-validating method in [15] relies on physical or analytical redundancy, which is more applicable to the data recovery of sensor arrays/networks. However, owing to the requirements of lightweight design, the structure of aerospace devices is too compact to mount multiple high-g accelerometers. As well, it is impractical to afford too many sensors because of the high price of each high-g accelerometer. Consequently, these redundancy-based methods are not adaptive to the high-g accelerometers used in pyroshock tests. Secondly, the low-rank completion method in [17] is also not suitable for the data recovery of high-g accelerometers because of the complicated frequency spectrums and drastically varying amplitudes. Lastly, the LS-SVM-based method in [18] is not applicable to the shock signals with high sampling rates.

In the past decades, data-driven methods have achieved great success in the health management of various systems [24-26]. These successes also inspired us to apply this powerful tool into the self-validation of high-g accelerometers. In our previous work, two deep neural networks (DNNs) were developed to recover shock signals of two forms of fault respectively [8, 9]. In addition, we also proposed an ensemble learning-based model to identify the fault types of high-g accelerometers by integrating several data-driven-based sub-classifiers [27]. However, these works cannot form an integral self-validation strategy to self-validate high-g accelerometers in pyroshock tests. On the other hand, in order to leverage the advantages of various data-driven algorithms adequately, hybrid strategies are usually proposed by combining two or more algorithms that are often exploited for solving more complicated problems [28]. Yu et al. [29] develop a hybrid strategy to accurately predict flight

trajectories of aircrafts by integrating the aircraft dynamics algorithm and the long short-term memory (LSTM) network. This method markedly improves the prediction accuracy and effectively decreases the computing costs compared with single machine learning-based methods. Neerukatti et al. [30] study a novel hybrid methodology to predict crack growths in aerospace structures via combining the energy release rate and a data-driven Gaussian process prognosis model. This methodology promotes the ability of conventional data-driven methods to predict the structural crack propagation under different loading conditions. Jahani et al. [31] propose a hybrid method to detect baseline shifts of motion artifact contamination in near-infrared spectroscopy data and correct them in parallel. This hybrid method integrates the spline interpolation algorithm and Savitzky–Golay filtering and provides improvements in multiple metrics with lower calculation costs. Inspired by these works, a hybrid strategy is proposed in this work to self-validate high-g accelerometers by integrating the ensemble kNNs and the DNN-based recovery method in our previous works [8, 9, 27].

In this work, considering the similarity between high-g accelerometers and pressure sensors [18] (both generate time series signals and are both widely applied in engineering mechanics), self-validating high-g accelerometers can also be decomposed into three parts: fault detection, fault identification, and data recovery. However, differing from the work in [18], which executes fault detection and fault diagnosis into two steps, we execute fault detection and fault diagnosis synchronously. With regard to the framework in Fig. 1, we construct the fault diagnosis module and the output data generation module with data-driven methods, mainly including ensemble kNNs and deep stacked autoencoders (DSAE)-based DNN. With this customized strategy, the core output parameters for self-validating sensors can be generated in our designed self-validating high-g accelerometer methodology. 1) The raw signals can be output directly. If the accelerometer is diagnosed as healthy, the raw signals are the measured shock signals of the healthy accelerometer, while if the accelerometer is diagnosed as faulty, the raw signals are the measured results of the faulty accelerometer. 2) The status of the measurements can be determined according to the fault detection results. In some application scenarios, especially the sensor arrays/networks conditions, the status can be divided into several types, such as secure status, clear status, blurred status, dazzled status, and blind status [11, 12]. In this work, the high-g accelerometers used in pyroshock tests usually work under standalone conditions, but not with sensor arrays/networks. Besides, in pyroshock testing scenarios, the damage of high-g accelerometer is usually a permanent failure. So, the status of the measurements in this work includes two classes, namely healthy status and faulty status, which can be determined at the fault detection step directly. 3) If the high-g accelerometer is detected as faulty, the specific fault type of this high-g accelerometer can be determined based on the fault diagnosis result. 4) The recovery signals can be the output, as the validated measurements, of the designed DSAE-based DNN, if the high-g accelerometer is diagnosed as faulty. 5) As for the validated uncertainty calculation with sensor faults, in this work, we propose a loss metric to measure the validated uncertainty of the proposed self-validating strategy under the corresponding fault type condition.

Lastly, a total of seven high-g accelerometers (with six faulty sensors and a healthy one) and an experimental apparatus are prepared for our case study. This experimental apparatus is customized by combining the vertical drop shock machine [32] and the dual mass shock amplifier (DMSA) [33], which can be used to simulate pyroshock environments accurately and generate shock data efficiently. The results of the case study demonstrate the validity of the proposed self-validation strategy, which can provide an effective self-validation solution to fault detection, fault identification, and data recovery of high-g accelerometers used in pyroshock tests. The main contributions of this work are highlighted as follows:

(1) A hybrid strategy is proposed to self-validate the high-g accelerometers by combining ensemble kNNs and DSAE-based DNN, which are capable of detecting the faulty measurements, identifying the fault types, and restoring desired data from the damaged measurements synchronously.

(2) To our best knowledge, this is the first attempt to introduce the self-validating sensor technique into the field of high-g accelerometers, contributing a new application in both the sensor self-validation field and the

aerospace pyroshock test field.

2. Hardware components

2.1 High-g accelerometers

As mentioned above, high-g accelerometers are the essential components in various shock test systems. Although diverse transduction mechanisms can all be applied for high-g sensing, the two most dominant sensors for measuring high-g shock are piezoresistive and piezoelectric accelerometers, widely used in both industry and academia [34]. Owing to that the piezoresistive and piezoelectric accelerometers cannot be compared directly, in this work, all the used accelerometers are the piezoelectric sensors. As shown in Fig. 2, a piezoelectric accelerometer mainly consists of package shell, mass block, piezoelectric element, base, preloaded screw, preloaded springs, and holder [35]. After mounted on the object to be tested, the base of the accelerometer will move together with the object with the same acceleration. In the meantime, the piezoelectric element will be subjected to the inertial force, which is opposite to the acceleration direction of the base. Subsequently, charges will be generated on the upper and lower surfaces of the piezoelectric element. Lastly, the acceleration of the testing object can be measured from the charges after amplification [27].

The piezoelectric accelerometer used for standard applications can be classified into three structural categories: shear, compressive, and flexion-based, of which the measuring ranges fall under 20000 g mostly [35]. In practice, the accelerometers' fault types typically include shearing of solder joints, chip crack, overstress of sensitive elements, etc. [35-37]. These faults will lead to erroneous measurements of the accelerometers, such as signal distortion, noise pollution, and baseline shifts [7-9, 38]. Note that, although the damage of the accelerometers' package shell is also a common failure occurs in practical high-g shock tests [39], this kind of failure can be diagnosed visually. So, the emphasis in this work is placed on the damage inner the package shell.
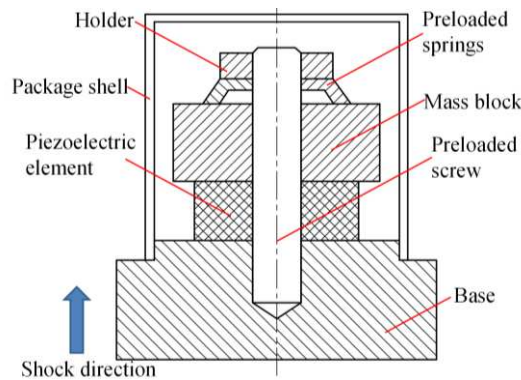

Fig. 2. Schematic diagram of piezoelectric accelerometer's structure.

2.2 Pyroshock test

Differing from common accelerometers, high-g accelerometers are mainly used for extreme mechanical environment testing, especially shock environment testing. However, different types of electronics would encounter different shock environments during their service-lives. For example, the shock environment of commercial electronics, such as mobile phones, laptop computers, audio players, and tablet PCs, is mainly caused by falls and transport, and, thusly, the shock test on the commercial electronics is usually termed as board-level drop test [8, 40]. On the other hand, for aerospace electronics, such as the camera in the space probe, various circuit boards in the satellite, and the onboard computer in the rocket, will undergo a series of pyroshock events throughout space missions, including the release of satellites, separation of launcher stages, the deployment of solar panels, the activation of impulse engines etc. Therefore, shock testing on aerospace electronics is referred to as pyroshock test, and the ground pyroshock test must be implemented to assess the reliability of the on-board

electronics before launching [4, 23].

In the drop test, the concern is focused on the main pulse of the shock signal, and the oscillation portion behind the main pulse is usually negligible. As an example shown in Fig. 3, two local characteristics, including the peak value and the pulse width, of the main pulse are extracted to evaluate the performance of the drop test [3]. In the pyroshock test, not only the peak value and pulse width, but also the frequency content should be both considered, and, thus, the main pulse and the oscillation portion should be taken into account synchronously.
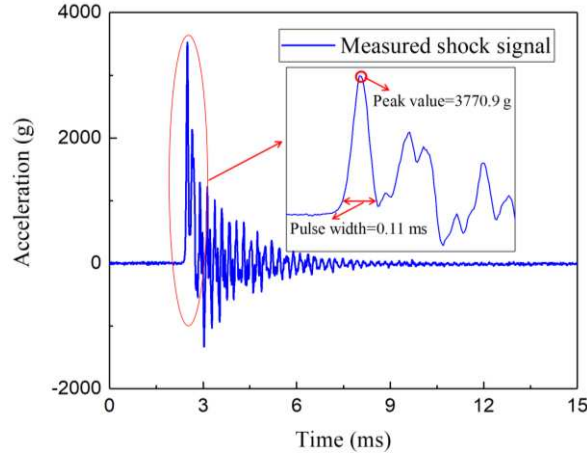
Fig. 3. The illustration of the peak value and the pulse width of a shock signal. The peak value is the maximum amplitude of the shock signal, and the pulse width is obtained at 10% the peak based on the JEDEC definition [3].

However, the pyroshock environment is usually too complex to elaborate a representative shock specification by a transient shock signal [41]. To solve this problem, the shock response spectrum (SRS) is developed as the shock specification to describe the shock severity in aerospace engineering [42]. The SRS curve is calculated by imposing the time-domain shock signals as the shock excitation into a series of single degree of freedom (SDOF) systems, and by plotting the system response w.r.t. increasing natural frequencies [42], i.e.,

$$S(f_i) = \max (A_i(t)). \tag{1}$$

In Eq.(1), $S(f_i)$ is the SRS value against the frequency f_i, and $A_i(t)$ is the time-domain response for the i-th SDOF system. The SRS curve can be computed quickly by converting the time-domain shock signal into a frequency-domain spectrum with the ramp invariant simulation method [43]. The SRS-based specification enables the comparison between different shock levels as well as different shock test facilities [41].

Generally, the pyroshock environment has less effect on the structure damage of products' package but will affect the function of the electronics inner the products' package to a great extent. The pyroshock environment can be divided into three categories: near-field, mid-field, and far-field based on the peak value and frequency content of the shock excitation [4]. Apart from the near-field pyroshock environment usually simulated with real pyrotechnic initiators, the mid- and far-field testing are usually performed via mechanical testing, such as the electromechanical shaker, pendular-hammer test benches, and free-fall shock test machines [44]. In our experiment, a space electronic device is tested with a high-g shock test system. As shown in Fig. 4, our testing system combines a vertical shock test machine and a DMSA. The DMSA is fixed on the drop table of a vertical shock test machine. The rebound springs hold the DMSA table above the DMSA base plate with a gap, and the DMSA table can move along the guide rods freely. To conduct the pyroshock test, the drop table and the DMSA will drop onto a rubber cushion mounted on the base rigidly following the guide rods by free-fall from a required height. Whereafter, the drop table will rebound upwards due to the elastic rubber cushion. At the same time, the DMSA table will keep up dropping and impact with the upward-moving DMSA base, which will generate a high-g shock response to simulate the pyroshock environment. A typical time-domain shock signal and its SRS curve are

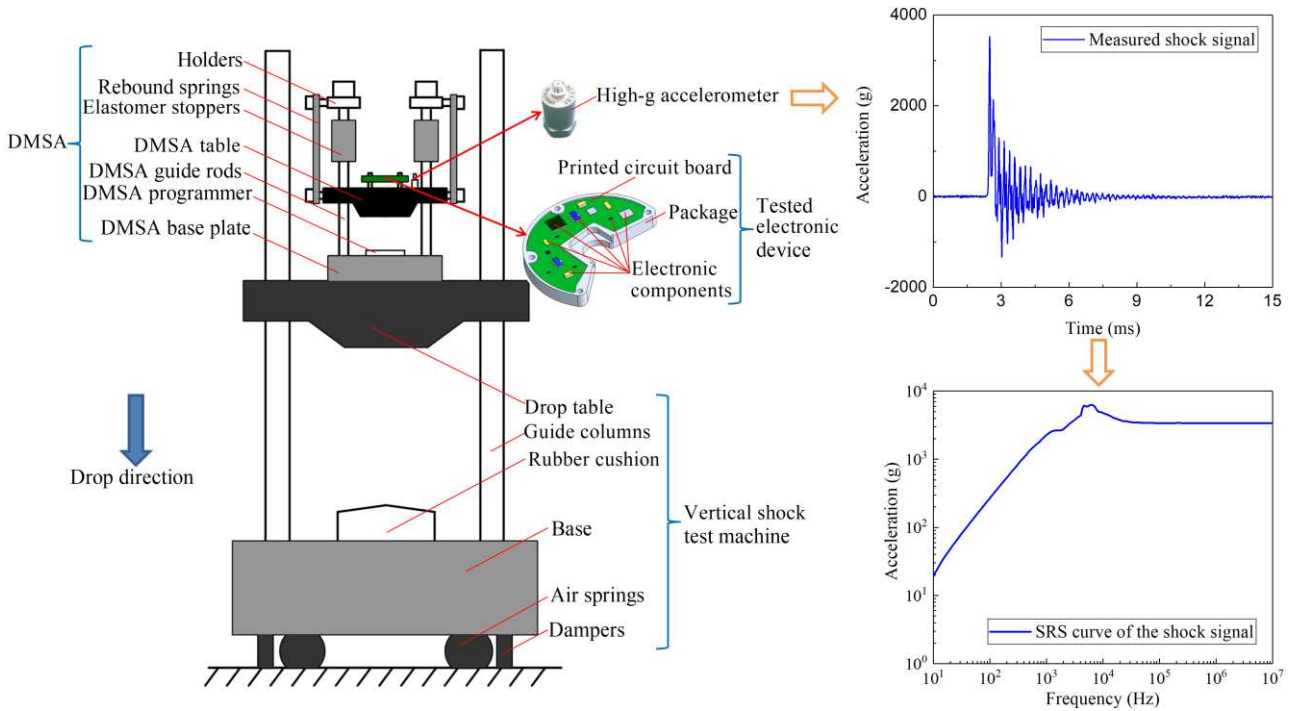presented in Fig. 4. The details of this shock test system can be referred to in [45].

Fig. 4. A typical free-fall pyroshock tester for testing aerospace electronics.

3. Software components

3.1 Ensemble kNNs

As a classical supervised learning approach, kNN will find k data points in the training set that has the smallest distance to a given testing data point, and the average of the labels to these k training samples will be treated as the prediction. Sensibly, the complexity of the kNN model is mainly dominated by the hyper-parameter k. Smaller k value corresponds to more complex model as the prediction from kNN could be more easily influenced by movements in the feature space. kNN is simple to be implemented, but our previous work [27] shows that standalone kNN classifier cannot achieve the best classification accuracy for faulty high-g accelerometers. A feasible strategy for improving the classification accuracy is integrating multiple individual kNN classifiers with ensemble learning [27, 46]. Inspired by these works, in this work, an ensemble learning model is constructed to identify the health conditions and diagnose the fault types for high-g accelerometers by applying Bagging algorithm to kNN classifiers with different k values. As shown in Fig. 5, the dataset is randomly sampled with replacement to generate the in-bag (IB) data, which are used to train the ensemble learning model; the unselected samples, as the out-of-bag (OOB) data, are used to validate and estimate the performance of the ensemble learning model. In this work, a total of eight sub-datasets are generated and used to train the base kNN classifiers (k=2, 3, …, 9) respectively. Finally, the classification results of the proposed ensemble learning model can be decided by aggregating the outputs of all the kNNs with the majority voting strategy.
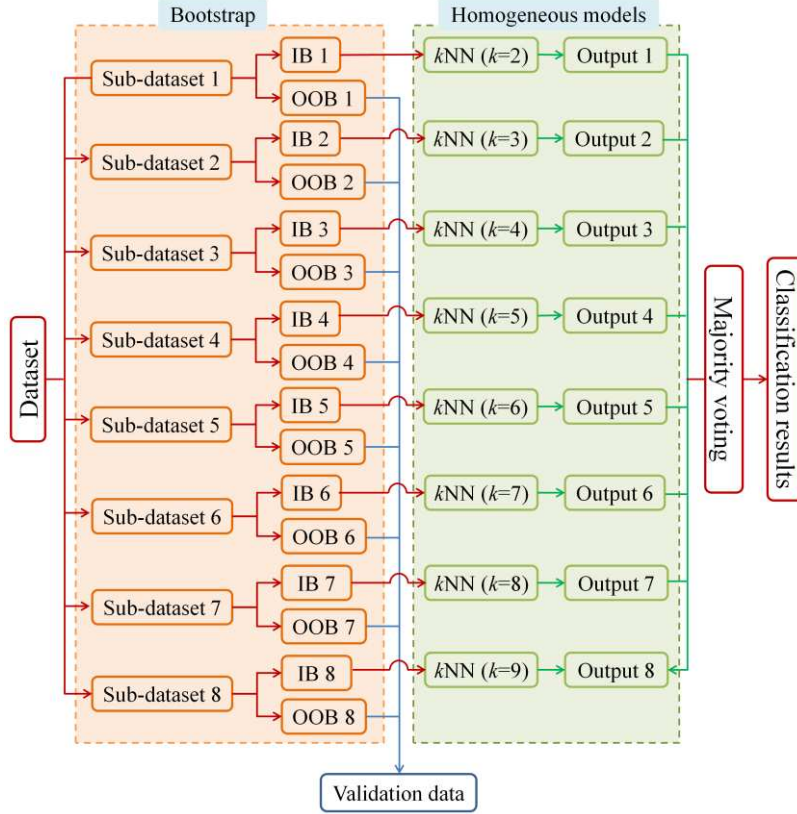
Fig. 5. Flowchart of the proposed ensemble learning model.

3.2 DSAE

DSAE is a popular deep learning method and has already demonstrated its capacity in mechanical damage diagnosis [47]. It is a hierarchical DNN made of multilayer autoencoders (AEs) [49]. As shown in Fig. 6(a), AE consists of an encoder and a decoder, of which the former encodes the input layer into an intermediate vector of lower dimensions, and the latter reconstructs the intermediate vector back to an output which has the same dimensions with the original input layer. Through minimizing the distance between the input and the output, AE will learn the features from the original data [49]. Given an unlabeled sample set $\{x^m\}_{m=1}^M$, randomly select a sample x^m which has the dimension of n as the input layer of the AE. The encoder maps x^m into an intermediate vector h^m with an encoding function f_θ, i.e.,

$$h^m = f_\theta(x^m) = s_f(Wx^m + b), \tag{2}$$

where s_f is the activation function of the encoder, and $\theta = \{W, b\}$ are the encoder's parameters. Subsequently, the intermediate vector h^m is projected back to the reconstruction vector $\hat{x}^m$ by the decoding function $g_{\theta'}$ as

$$\hat{x}^m = g_{\theta'}(h^m) = s_g(W'h^m + b'), \tag{3}$$

in which s_g is the activation function of the decoder, and $\theta' = \{W', b'\}$ are the decoder's parameters. The training of AE is done by minimizing the distance between $\hat{x}^m$ and x^m, which can be expressed as:

$$\min L(\hat{x}^m, x^m) = \min \frac{1}{M} \sum_{m=1}^M \left(\frac{1}{2} \|\hat{x}^m - x^m\|^2 \right), \tag{4}$$

where $L(\hat{x}^m, x^m)$ is the reconstruction error (Euclidean distance is used as an example in Eq.(4)), and M is the number of all the samples. If $\hat{x}^m$ is sufficiently similar to x^m, it means that h^m extracts most of the features comprised in the input layer.

DSAE expands the concept of AE by including multiple hidden layers between the input and the output. As depicted in Fig. 6(b), DSAE is obtained by linking AEs layer by layer. The training process of DSAE usually involves two steps: pre-training and fine-tuning [47]. The pre-training is an unsupervised learning process, in

which the first AE maps the input layer into the first hidden layer by minimizing the reconstruction error. After training the first AE, the parameters $\{W_1, b_1\}$ can be obtained. Then, eliminating the output layer of the first AE, the hidden layer of the first AE is served as the next input layer for the second AE. Similarly, the second AE is trained and the parameters $\{W_2, b_2\}$ can be obtained. In the similar fashion, the entire DSAE can be pre-trained layer by layer until finish training the last AE, and all the parameters $\{W_i, b_i\}_{i=1,2,\ldots,n}$ can be acquired preliminarily. The fine-tuning is a supervised learning process followed with the pre-training. In this process, the labeled output layer is set on the rightmost of the DSAE and the training parameters are fine-tuned. Specifically, setting the pre-trained parameters $\{W_i, b_i\}_{i=1,2,\ldots,n}$ as the initialization parameters of each hidden layer and randomly initializing the training parameters of the output layer $\{W_{n+1}, b_{n+1}\}$, the whole DSAE network can be trained by the back propagation algorithm. The optimal fine-tuned parameters $\{W_i', b_i'\}_{i=1,2,\ldots,n+1}$ can be finally obtained by minimizing the predicting error, i.e.,

$$\min L(y^m, f_\theta'(h_n^m)) = \min \frac{1}{M} \sum_{m=1}^{M} \left(\frac{1}{2} \|y^m - f_\theta'(h_n^m)\|^2 \right), \tag{5}$$

in which f_θ' is the target classifier (or other neural networks); $f_\theta'(h_n^m)$ is the predicted output; and y^m is the labeled output.

Compared with AE, DSAE can extract more complex and abstract features from the input data. Therefore, DSAE tends to have better modeling performance and can approximate more complicated functions [47].
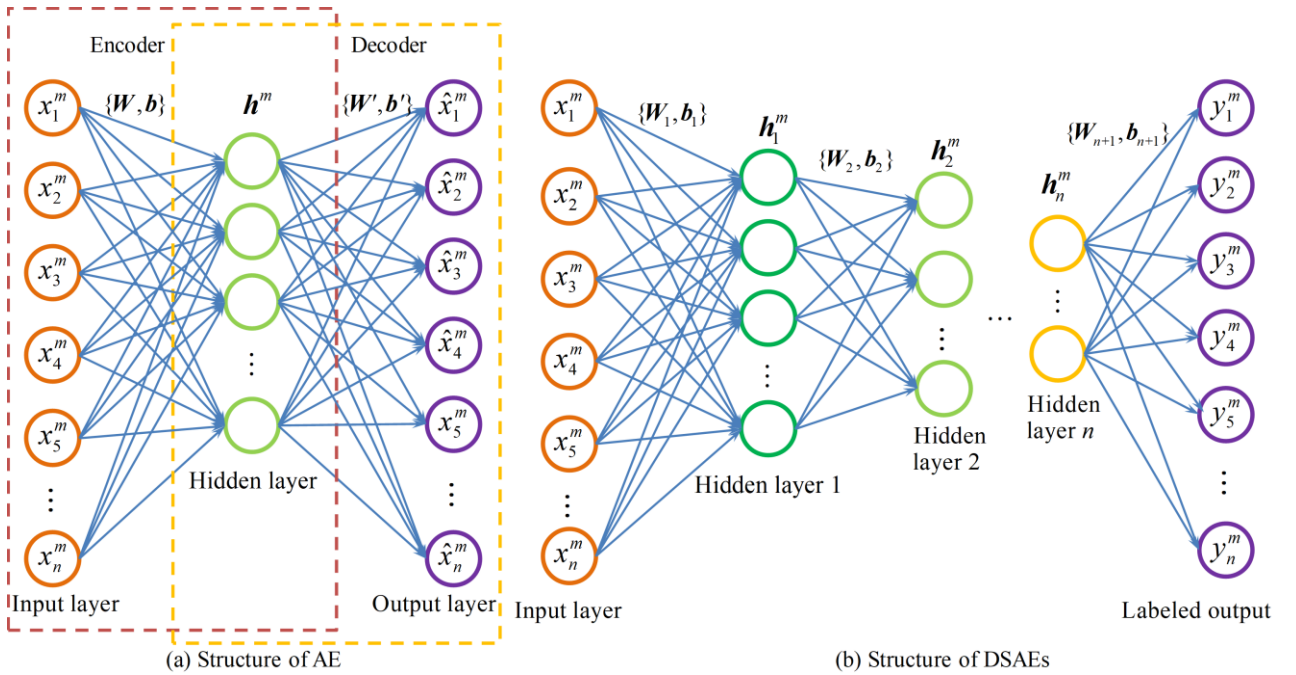

Fig. 6. Structures of AE and DSAE.

3.3 The motivation of utilizing DSAE method

The inspiration of utilizing DSAE method to implement the data recovery of faulty shock signals derives from Jelinek's work [48] in the speech recognition domain. In his work, as shown in Fig. 7, the speech recognition process is equivalent to a communication model. Jelinek considered human's brain as an information processing mechanism, in which speaker's voicing is an encoding process, and the listener's understanding is a decoding process.
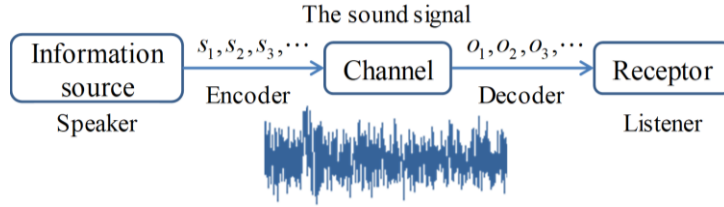
Fig. 7. The communication model of speech recognition. In this figure, s_1, s_2, s_3, … denotes the signal generated by the infroamtoin source; o_1, o_2, o_3, … is the signal received by the receptor.

Similar to humans' ears to perceive sound signals and decode the perceived sound signals into comprehensible language information, the high-g accelerometer perceives the real shock signals in the physical world and decodes the perceived shock signals into acceleration-time series that can be understood by humans. As shown in Fig. 8, the true shock signal can be generated with a certain encoding rule, when a real shock event happens in the physical world; then, the true shock signal will be transmitted to the accelerometer through the bolt or other fasteners (equivalent to the channel of communication system); finally, the job of the accelerometer is to decode the transmitted shock signal into the acceleration-time series. Analogously, the encoding rule is the classical mechanics, the designing principle of accelerometers, and Fig. 2 shows the decoding rule, which is a process of inverse classical mechanics.

In Fig. 8, encoder 1 and encoder 2 are identical, and the reason that the healthy accelerometer and the faulty accelerometer output different shock signals is the difference between decoder 1 and decoder 2. The calibration experiment for accelerometers can prove that decoder 1 is correct. Traditionally, the most conservative solution to improve the performance of decoder 2 is to fix the hardware deficiency of the faulty accelerometer. In this work, the hypothesis is that, if the mapping between the decoded signals of the healthy accelerometer and the faulty accelerometer can be constructed, it is equivalent to correcting decoder 2.
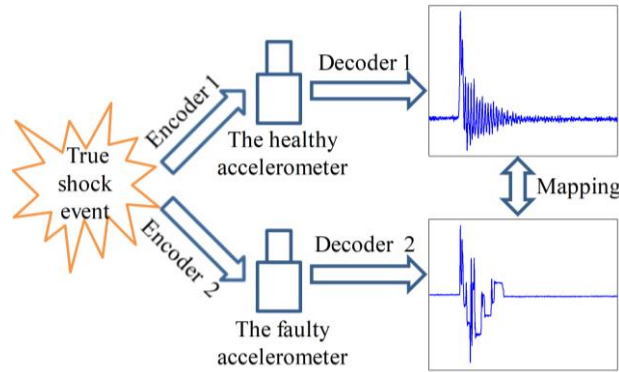

Fig. 8. The working principle of high-g accelerometers based on the communication model.

If decoder 2 is proved to be correct, the shock signal is then considered a correct description for the true shock event. As shown in Fig. 9(a), substituting the true shock event with the shock signals measured by the healthy accelerometer, and assuming that there is a virtual accelerometer, which can decode the shock signals measured by the healthy accelerometer into the measuring result of the faulty accelerometer, the communication model between the decoded signals of the healthy accelerometer and the faulty accelerometer can be constructed. Then, as shown in Fig. 9(b), reversing the communication model in Fig. 9(a), i.e. setting the shock signal measured from the faulty accelerometer as the input and the shock signal measured from the healthy accelerometer as the output, the mapping of recovering the faulty shock signals can be constructed if encoder 3' and decoder 3' can be attained. Additionally, it can be found that the structures of DSAE in Fig. 6 is similar to the mapping structure in Fig. 9(b) if equalizing the virtual accelerometer as an abstract feature vector extracted by the DSAE network. Thusly, in this work, a DSAE-based method is adopted to construct the DNN to recover the faulty shock signals.
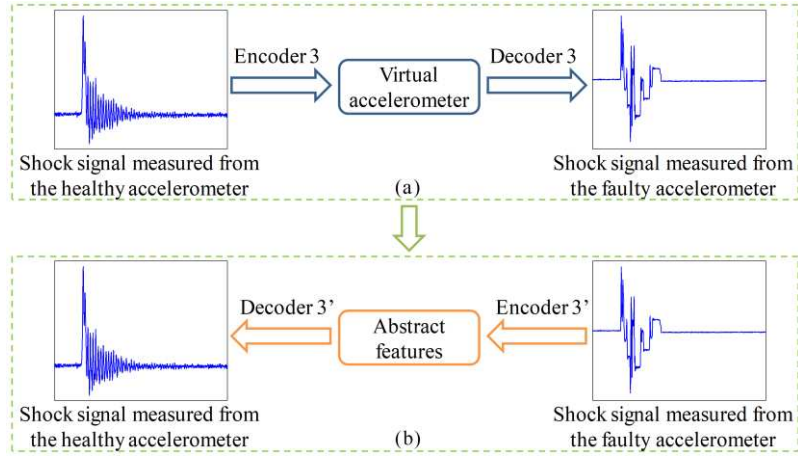
10

Fig. 9. Schematic diagram of the data recovery method for high-g accelerometers based on the communication model.

3.4 Strategy of the proposed self-validation method

In this work, a data-driven-based strategy is proposed by combining the ensemble kNNs and DSAE-based DNN to self-validate high-g accelerometers used in pyroshock tests. The inspiration of this strategy stems from the medical consultation: The doctor shall firstly detect the patient is healthy or not; then, the doctor shall identify the patient's specific symptoms, if the patient is confirmed to be in an unhealthy status; lastly, the doctor shall provide specific treatments to cure the unhealthy patient as much as possible. Analogically, the proposed strategy for self-validating high-g accelerometers should be able to fulfill these steps as well. To this end, also, a data-driven-based strategy for self-validating accelerometers is designed and shown in Fig. 10.
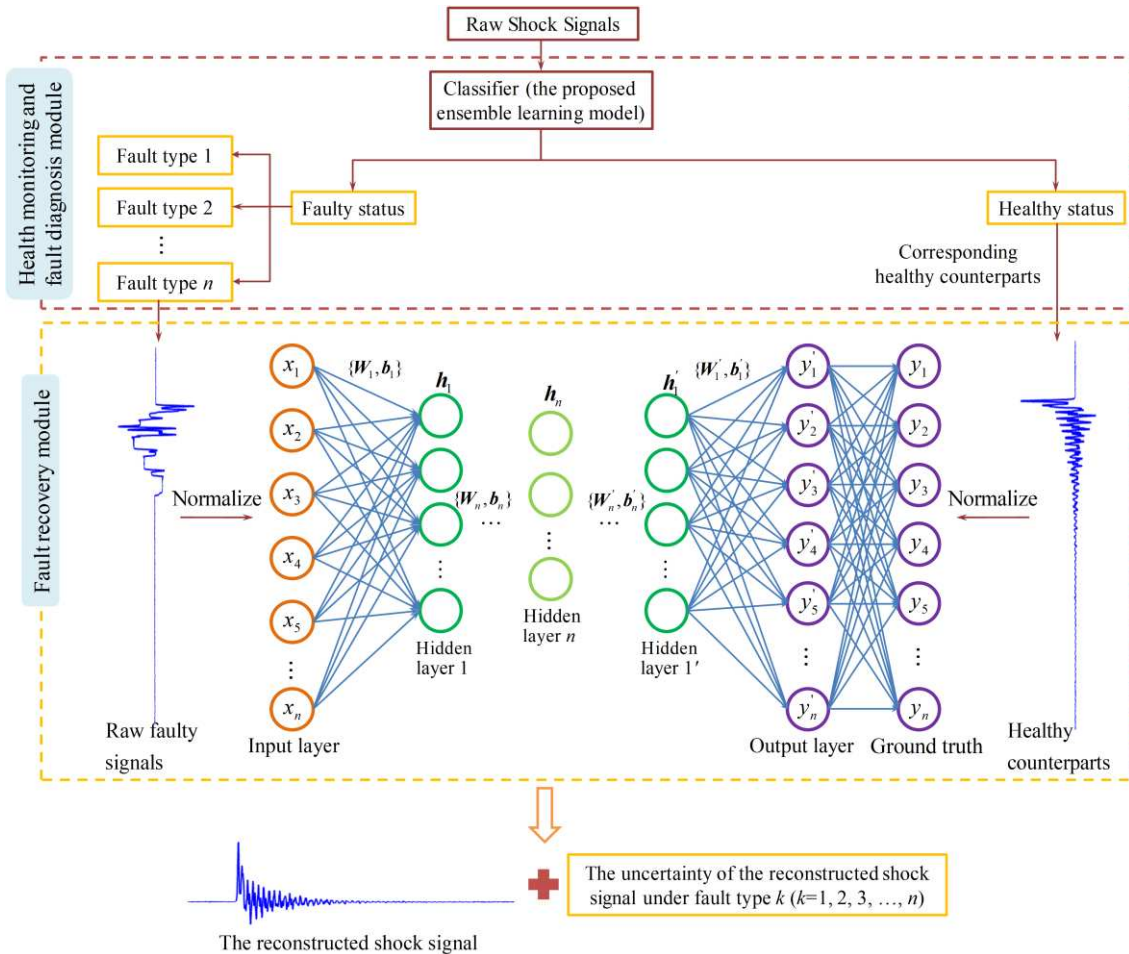
Fig. 10. The strategy of self-validating high-g accelerometers in this work.

11

In Fig. 10, the measured shock signals are firstly input into the classifier (i.e. the ensemble kNNs), which can detect if the input signals measured are healthy or faulty and subsequently can classify the faulty signals into specific fault types. Differing from traditional DSAE-based fault diagnosis strategies that the healthy data are merely set as the reference [53-55], in our work, the healthy shock signals are utilized as not only the reference to identify the health conditions and fault types, but also the ground truth to be recovered from the faulty signals. Firstly, as shown in Fig. 10, the diagnosed faulty shock signals (with one kind of fault type) are normalized as the input layer and a DSAE-based network with $2(n+1)$ layers can be pre-trained with greedy layer-wise strategy. Then, normalize the healthy counterparts as the ground truth and the whole network can be finally trained end-to-end by minimizing the reconstruction error between the output-layer vector and the ground truth. Lastly, the recovered shock signal and its uncertainty under the corresponding fault type calculated with the loss metric can be output synchronously. The whole data-driven strategy for the self-validation system can be built up by hybridizing all the classifiers and networks.

The details of our designed DSAE-based DNN are demonstrated in Fig. 11, which are also used to correct a bunch of heavy noise-polluted shock signals in our previous work [9]. Intuitively, the designed DSAE includes three parts: encoder, decoder, and a customized peak prediction network (PPN). The input layer firstly normalizes the diagnosed faulty shock signals (with one kind of fault type) x_r before feeding into the DSAE, and the normalized healthy signals is used as the ground truth to guide the training of the entire network in an end-to-end fashion. Specifically, the encoder first encodes the normalized faulty shock signal $\tilde{x}_r$ into a 256-dimensional vector $\tilde{x}_n$, and the decoder reconstructs back it to a 2000-dimensional vector $\tilde{y}_{out}$:

$$\begin{cases} \tilde{x}_n = f_{enc}(\tilde{x}_r; \theta_1) \\ \tilde{y}_{out} = f_{dec}(\tilde{x}_n; \theta_2) \end{cases}, \tag{6}$$

where f_{enc} and f_{dec} are the encoder and decoder functions respectively; θ_1 and θ_2 are the network parameters for the encoder and decoder separately.

As the peak value is crucial for shock signals, the PPN is specially designed to preserve the peak information in the recovery shock signals with high fidelity. In our previous work [9], the effect of PPN is verified effectively regarding to the fault type of noise-pollution. The global feature vector $\tilde{x}_n$ is further encoded into 8 dimensions and then connected with the peak value of the input shock signal. Then, the feature vector $\tilde{x}_n$ and the peak value $\tilde{p}_r$ are fed into the PPN simultaneously:

$$\begin{cases} \tilde{p}_{err} = f_{PPN}(\tilde{p}_r, \tilde{x}_n; \varphi) \\ \tilde{p}_{out} = \tilde{p}_r + \tilde{p}_{err} \end{cases}, \tag{7}$$

where f_{PPN} is the PPN function, and φ is the network parameter of the PPN; $\tilde{p}_{err}$ is the estimated error between the accurate peak value and the input peak value, which will be added back to $\tilde{p}_r$ to predict the correct peak value $\tilde{p}_{out}$. The reconstructed signal y_{pred} can be attained by unnormalizing $(\tilde{y}_{out} + \tilde{p}_{out})$.

We further use y_{ref} and $\tilde{y}_{ref}$ to denote the unnormalized and normalized healthy signal. All networks can be trained by minimizing the following loss functions:

$$\begin{cases} L_{wave}(\theta_1, \theta_2) = \|\tilde{y}_{out} - \tilde{y}_{ref}\|_2 + \|\tilde{y}_{out} - \tilde{y}_{ref}\|_\infty \\ L_{peak}(\varphi) = \|\tilde{p}_{out} - \tilde{p}_{ref}\|_2 \end{cases}, \tag{8}$$

where L_{wave} is the loss to regulate the normalized shock signal, which is parametrized with θ_1 and θ_2; L_{peak} is the loss to regulate the peak value of the faulty signal, which is parametrized with φ (where θ_1 is kept frozen). We expect our trained network can transform the raw signal (i.e. x_r) produced by the faulty accelerometers to a reconstructed signal (i.e. y_{pred}) similar to what the healthy accelerometer will produce.
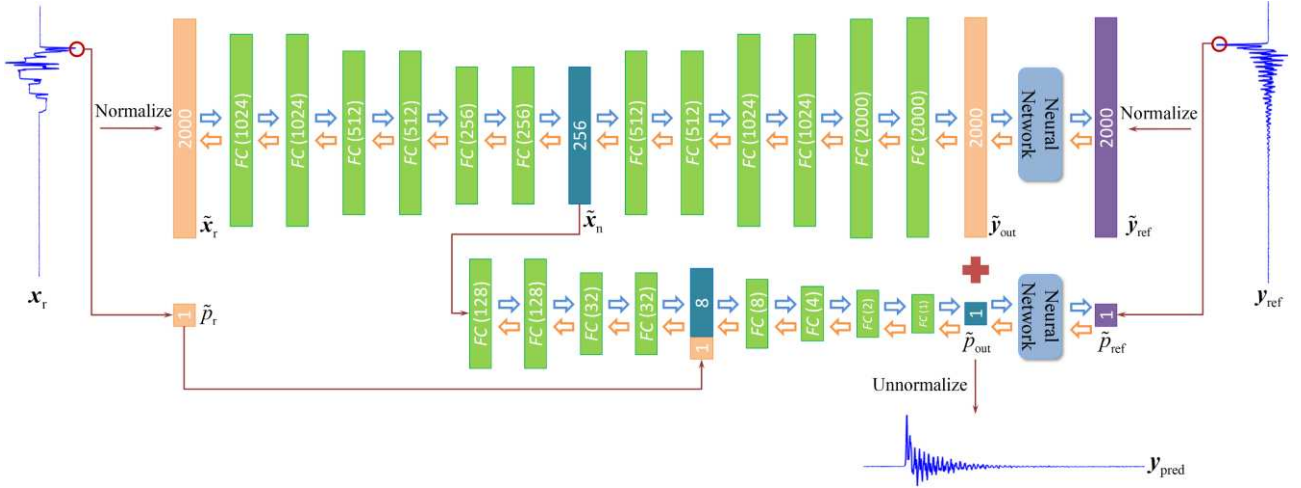
Fig. 11. The architecture of the designed DSAE.

4. Experiments

As a case study, this section validates the proposed self-validation system on an industrial pyroshock test system. This section is divided into three parts: 1) the experimental system and the data acquisition process are introduced in detail; 2) the performance metrics for assessing the proposed model are presented; and 3) the experimental results are demonstrated.

4.1 Data acquisition

In this work, based on the working principle of the pyroshock tester introduced in Section 2.2 and the experimental configuration diagramed in Fig. 4, a pyroshock test system, as shown in Fig. 12, is designed to establish the pyroshock environments for data acquisition. In this system, different shock levels can be produced by adjusting the falling heights of the drop table, and a dataset comprising vast shock signals measured under different shock levels can be collected by operating the experimental system repeatedly. All these shock signals are measured with a healthy accelerometer and six faulty accelerometers synchronously, which guarantees that all the accelerometers work at the same shock level during each shock test. All the measured shock data are recorded by the data acquisition and processing system and can be indicated on the monitor directly, prior to which the raw shock signals should be amplified by the charge amplifier. The sampling rate of the data acquisition system is 200 kHz, which ensures the high fidelity of all the measured shock data.

The relative calibration method is applied in the data acquisition process. However, as shown in Fig. 13, differing from the traditional "back-to-back" mounting style, a "central-symmetry" mounting style is employed in this work to implement the synchronous measurements of multi-accelerometers at the same shock level. Since each accelerometer is symmetrically distributed in the center of the DMSA table, the same shock response measured by each accelerometer can be implemented as long as the table is with high uniformity and the shock excitation is acted on the center of the table. Comparing with the "back-to-back" mounting style, the "central-symmetry" mounting style possesses higher calibration efficiency, with which four accelerometers can measure the same shock response simultaneously at one time. Although the "back-to-back" mounting style is more conducive to guarantee calibrating accuracy, the calibration efficiency should be considered preferentially because the accumulation of vast samples for data-driven methods. However, it does not mean that the calibration accuracy is not important in this work.
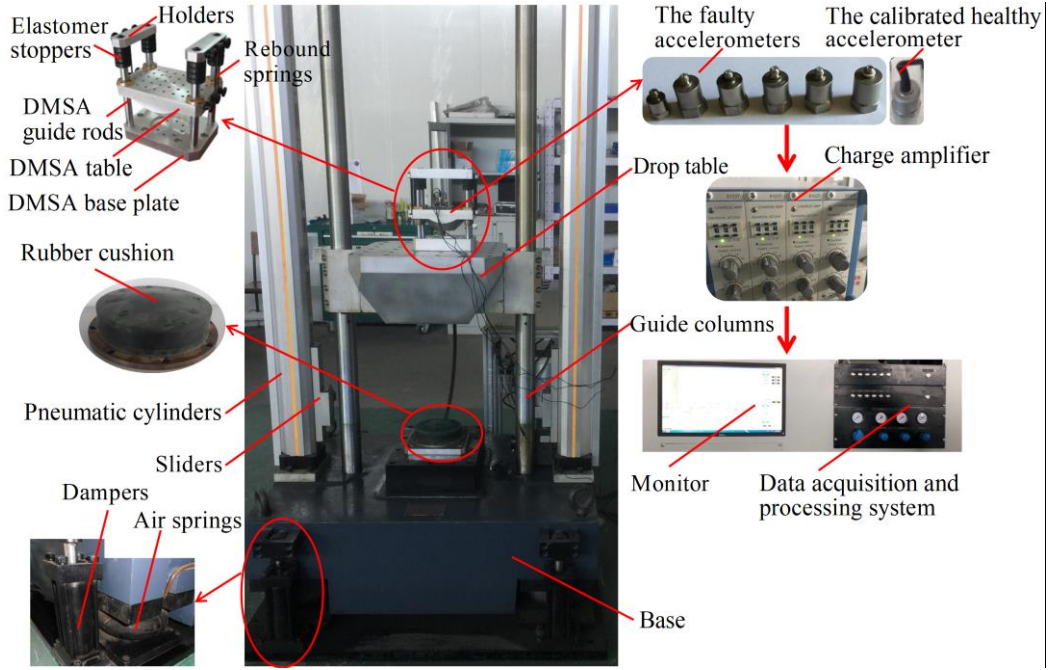
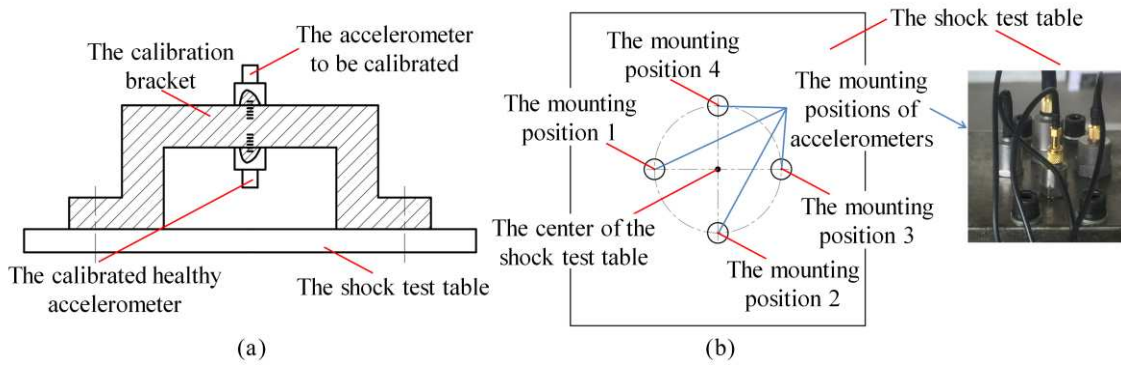Fig. 12. The established pyroshock tester. This tester is adapted from [27].

(a)　　　　　　　　　　　　(b)

Fig. 13. The mounting styles of the relative calibration method. In this figure, (a) is the "back-to-back" mounting style; (b) is the "central-symmetry" mounting style.

In order to verify the calibration accuracy of the "central-symmetry" mounting style, the calibrated healthy accelerometer was mounted at four positions respectively to measure the shock response at the same shock level. The shock signals measured by the calibrated healthy accelerometer are demonstrated in Fig. 14. Without loss of generality, set the shock signal measured at the mounting position 1 as the reference signal, and, then, the shock signals measured at mounting position 2, 3, 4 can be compared with the reference signal respectively. In this work, the determination coefficient is used to measure the approximation between two shock signals measured at mounting position 1 and another mounting positions quantitatively. The calculation formula can be expressed as follows:

$$R_{1-j}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^j - y_i^1)^2}{\sum_{i=1}^{n}(y_i^j - \bar{y}^1)^2},$$
(9)

where j=2, 3, 4; R_{1-j}^2 is the determination coefficient between the shock signals measured by the calibrated healthy accelerometer at mounting position 1 and mounting position j; n is the number of sampling points of the measured shock signal; y_i^j is the value at i-th point of the shock signal measured at mounting position j; y_i^1 is the value at i-th point of the shock signal measured at mounting position 1; and $\bar{y}^1$ is the average value of all the points in the shock signal measured at mounting position 1. Obviously, the closer R_{1-j}^2 is to 1, the closer the shock signal measured at mounting position 1 is to that measured at mounting position j. The calculated results are

14

recorded in Table 1. It can be seen that the calibration results in Table 1 are very close to 1, meaning that the shock response measured at mounting positions 2, 3, 4 are very close to the shock response measured at mounting position 1. Therefore, it can be verified that the "central-symmetry" mounting style adopted in this work is feasible to carry out the data acquisition of high-g accelerometers.
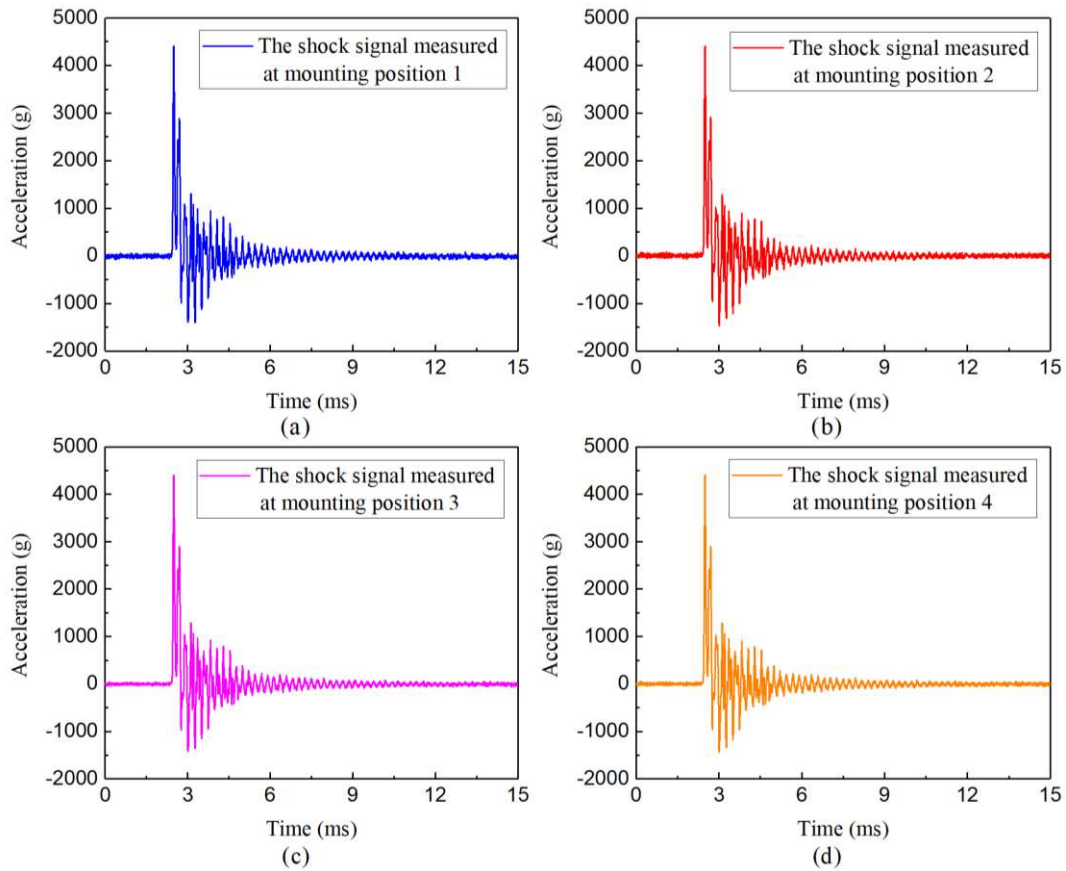

Fig. 14. The shock signals measured by the calibrated healthy accelerometer at four different mounting positions.

Table 1. The results of R^2_{1-j}.

R^2_{1-2}	R^2_{1-3}	R^2_{1-4}
0.985	0.996	0.992

As shown in Fig. 15, the seven collected high-g accelerometers are numbered in serial. Among them, accelerometer 1 is the calibrated healthy accelerometer, and accelerometers 2, 3, …, 7 are the faulty accelerometers collected from actual shock test tasks. As there are only four mounting positions with the "central-symmetry" mounting style, all the collected accelerometers should be divided into two batches to conduct shock test experiments. In the first batch of shock test experiment, the used accelerometers include accelerometers 1, 2, 3, 4, while, in the second batch of shock test experiment, the used accelerometers incorporate accelerometers 1, 5, 6, 7. Likewise, without loss of generality, accelerometer 1 is always mounted on mounting position 1.
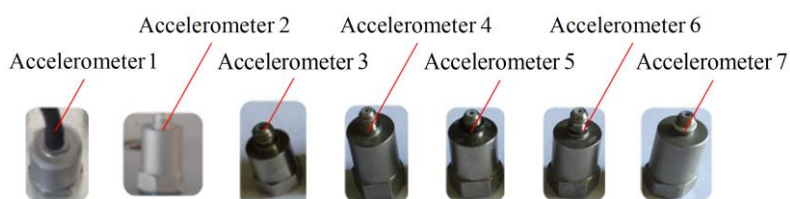

Fig. 15. The seven collected high-g accelerometers in this work.

15

In this work, a dataset contains 3004 sets of shock signals was collected by implementing the shock tests 751 times at different shock levels. The detailed description of the dataset and the fault types of each accelerometer are summarized in Table 2. The typical shock signals measured from each accelerometer are displayed in Fig. 16. More details of this dataset can be referred in [27]. Owing to the low repeatability of faulty accelerometers, it can be observed from Fig. 16 that some faulty accelerometers with one fault type could generate mutable faulty shock waveforms, and different faulty accelerometers could output similar faulty waveforms. This phenomenon raises the difficult level of accurately diagnosing the fault types of the accelerometers from its erroneous readings as well as recovering the faulty signals. Additionally, it should be noted that all these faulty shock signals are measured with the accelerometers used in this work. In practice, although the accelerometers have the same fault type, different signal waveforms can still be generated under the conditions of different shock environments and/or sensor types.

Table 2. The description for the collected dataset of shock signals.

| Accelerometer label	Fault types	Number of samples	Fault type label
Accelerometer 1	No fault	751	Type 1
Accelerometer 2	Overstress of the sensitive element	510	Type 2
Accelerometer 3	Damage of the sensitive element	510	Type 3
Accelerometer 4	Mounting base loosening	510	Type 4
Accelerometer 5	Sensitive element loosening	241	Type 5
Accelerometer 6	Mass block loosening	241	Type 6
Accelerometer 7	Low resonant response	241	Type 7
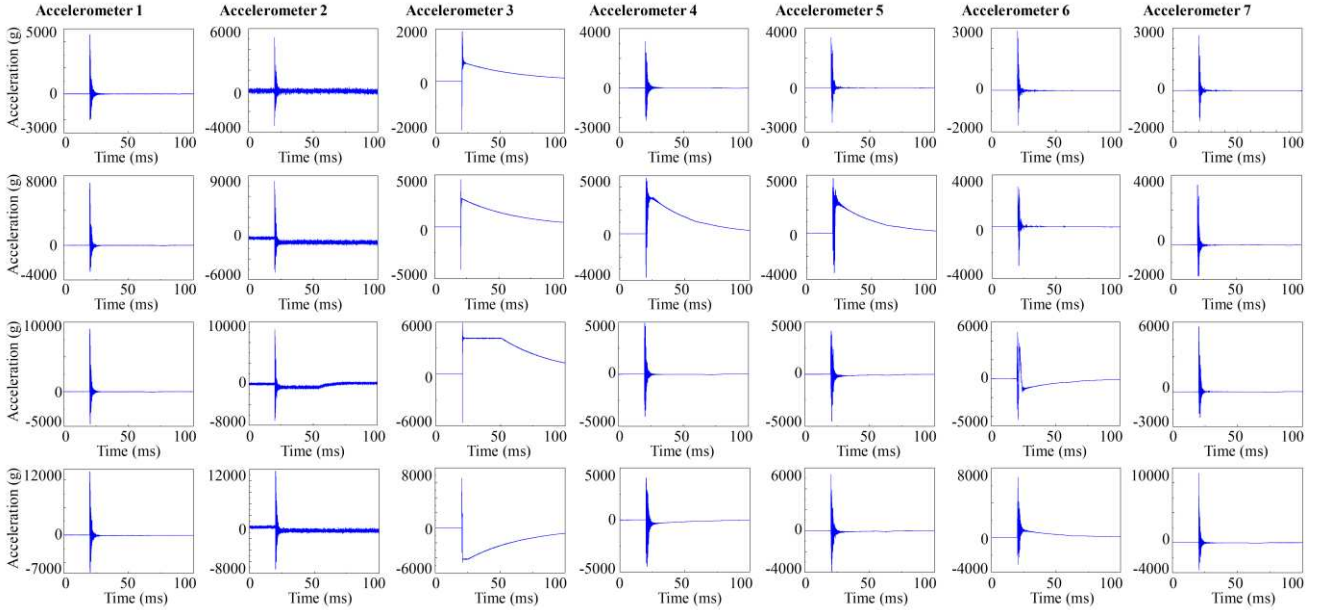
Fig. 16. Typical readings of each accelerometer.

4.2 Performance metrics

Based on the definition of self-validation sensors, the metrics for evaluating the performance of the proposed methodology should be formulated mainly in terms of the fault detection, fault diagnosis, fault recovery, and uncertainty on the recovery result.

Since that fault detection and fault diagnosis are both classification tasks (the former is a binary classification and the latter is a multi-class classification) essentially, the accuracy (A) and F_1-score (F_1) [50, 51] are adopted for evaluating the performance of the proposed fault detection and fault diagnosis methods. The formulations of A and

F_1 can be defined as

$$
\begin{cases}
A = \dfrac{TP+TN}{TP+TN+FN+FP}, \\
F_1 = \dfrac{2TP}{2TP+FP+FN}
\end{cases}
\tag{10}
$$

where TP is the number of samples correctly classified in the positive category; TN is the number of samples correctly classified in the negative category; FN is the number of samples misclassified in the negative category; and FP is the number of samples misclassified in the positive category. In Eq. (10), A illustrates the overall performance of the classification results, and F_1 reflects the classification performance of each category. Higher A and F_1 indicate better performance of fault detection and fault diagnosis respectively, where the maximum values of A and F_1 are both 100% in this work. Meanwhile, the fault detection and fault diagnosis results can be intuitively demonstrated with the confusion matrix, which is also one of the commonly used visualization methods for classification problems [50-54].

However, fault self-recovery is naturally a data reconstruction problem, and there is no generally accepted metric for estimating the performance of data recovery. In our previous work [8], four metrics, including the relative error of peak values, difference of time-domain waveforms, determination coefficient of time-domain waveforms, and determination coefficient of SRS curves, were proposed to quantify the signal recovery performance by considering both the board-level drop shock test and pyroshock test synchronously. Among these metrics, the relative error of peak values describes the ability of recovering maximal shock responses from the faulty signals, which is a key index to evaluate the reliability of products during board-level shock tests; the difference of time-domain waveforms describes the local accuracy of every point in the reconstructed shock signal; the determination coefficient of time-domain waveforms indicates the global accuracy of the overall reconstructed shock signal, which is also a precondition to calculate the corresponding SRS curve; and the determination coefficient of SRS curves reflects the accuracy of the reconstructed shock signal for assessing the reliability of products during pyroshock tests. Considering that the background of this work focuses on the self-validation of high-g accelerometers used in pyroshock tests of aerospace electronics, the SRS-based shock test standards should be mainly concentrated on and the board-level shock test can be ignored. Thusly, in this work, the latter two metrics, i.e. the determination coefficient of time-domain waveforms and the determination coefficient of SRS curves, are adopted as the performance metrics to estimate signal recovery results. These two metrics can be respectively defined as

$$
\begin{cases}
R_{\mathrm{T}}^2 = 1 - \dfrac{\Sigma_1^n\left(y_i^{\mathrm{Rec}}-y_i^{\mathrm{Ref}}\right)^2}{\Sigma_1^n\left(y_i^{\mathrm{Rec}}-\bar{y}^{\mathrm{Ref}}\right)^2} \\
R_{\mathrm{S}}^2 = 1 - \dfrac{\Sigma_1^m\left(s_i^{\mathrm{Rec}}-s_i^{\mathrm{Ref}}\right)^2}{\Sigma_1^m\left(s_i^{\mathrm{Rec}}-\bar{s}^{\mathrm{Ref}}\right)^2}
\end{cases}
\tag{11}
$$

In Eq. (11), R_{T}^2 is the determination coefficient between the recovery shock signal and the reference shock signal in the time domain; R_{S}^2 is the determination coefficient between the SRS curve of the recovery shock signal and the SRS curve of the reference shock signal; $\boldsymbol{y}^{\mathrm{Rec}}$ and $\boldsymbol{y}^{\mathrm{Ref}}$ represent the recovered and reference shock signal series respectively; $\bar{y}^{\mathrm{Ref}}$ is the mean of $\boldsymbol{y}^{\mathrm{Ref}}$; y_i^{Rep} and y_i^{Ref} are the value of the i-th point in $\boldsymbol{y}^{\mathrm{Rec}}$ and $\boldsymbol{y}^{\mathrm{Ref}}$ respectively; similarly, $\boldsymbol{s}^{\mathrm{Rec}}$ and $\boldsymbol{s}^{\mathrm{Ref}}$ denote the SRS series of the recovered and reference shock signals respectively; s_i^{Rec} and s_i^{Ref} are the value of the i-th point in $\boldsymbol{s}^{\mathrm{Rec}}$ and $\boldsymbol{s}^{\mathrm{Ref}}$ respectively; n and m are the dimentionalities of the shock signal series and the SRS series respectively. The closer to 1 the R_{T}^2 or R_{S}^2 is, the closer the recovery shock signal is to the reference shock signal in the time domain and the frequency domain respectively, hence, higher recovery performance. Actually, the determination coefficient is a scientific indictor to measure the degree of approximation of two curves specifically [52]. In this work, Eq. (9) and Eq. (11) are both designed based on the definition of the determination coefficient.

As for the metric of the validated uncertainty, there is unfortunately not a unified and generally accepted

methodology. Usually, the validated uncertainty with sensor fault can be described as

$$\delta_k = \delta_0 + \delta_{\text{add}}, \tag{12}$$

where δ_k is the validated uncertainty to be calculated at k-th fault type, and it can be found from Table 2 that k=1, 2, 3, …, 7; δ_0 is the uncertainty value of the sensor at healthy status; δ_{add} is usually determined by experts' experience according to different fault modes [18]. Sensibly, this method would import a lot of human interference.

As the validated uncertainty is used to measure the effect of the fault type on the raw measurements, without loss of generality, it can be assumed that δ_0=0. On the other hand, it can be found in our previous work [9] that one of the loss metrics, as expressed in Eq. (13), reflects the recovery capacity of the trained DNN.

$$E_k = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{N}\frac{\left|y_{mi}^{\text{Rec}} - y_{mi}^{\text{Ref}}\right|}{f_{\max}(y_m^{\text{Ref}})}. \tag{13}$$

In Eq. (13), E_k presents the recovery performance of the proposed data recovery method for the overall waveform of the shock signal under the k-th fault type; M is the number of testing samples under the k-th fault type; N is the dimensions of the shock signal; $f_{\max}$ is the function to calculate the maximum value of the shock signal series; y_m^{Rec} and y_m^{Ref} are the recovery shock signal and reference shock signal corresponding to the m-th sample in the testing set; and y_{mi}^{Rec} and y_{mi}^{Ref} are the values of the i-th point in y_m^{Rec} and y_m^{Ref} respectively. In this work, we utilize this loss metric to measure the uncertainty of recovery results under different fault types, i.e.

$$\delta_k = \delta_0 + E_k. \tag{14}$$

Obviously, the lower E_k is, the better the recovery performance is, hence a lower uncertainty (i.e. more trustworthy validated measurements). If E_k is 0, the uncertainty of the measurements will be the lowest (actually, it is the measuring results of the healthy accelerometer).

4.3 Results and analysis

In the collected shock dataset, about 80% of the samples for each accelerometer were randomly extracted to form the training set, and approximate 20% were used for testing.

To analyze the performance of fault detection and fault diagnosis, the confusion matrix of the proposed classification methods is visualized in Fig. 17. The confusion matrix gives the correctly classified samples and misclassified samples for each individual category. In this confusion matrix, the columns and rows represent the true labels and the predicted labels respectively, and the diagonal blocks show the number of samples that are correctly classified. Additionally, as shown in Table 3, the fault detection and fault diagnosis performance of all the proposed methodology are quantified with two indicators: the accuracy and F_1-score. Note that, in this work, the fault detection and fault diagnosis are executed synchronously by designing ensemble kNNs, but it does not mean the fault detection and fault diagnosis must be implemented together. In practice, as the flexibility of the ensemble kNNs, users can easily adjust the architecture of the ensemble kNNs based on their practical requirements: fault detection only, fault diagnosis only, or both fault detection and diagnosis.

Table 3. Accuracy and F_1-score of the fault detection and fault diagnosis results.

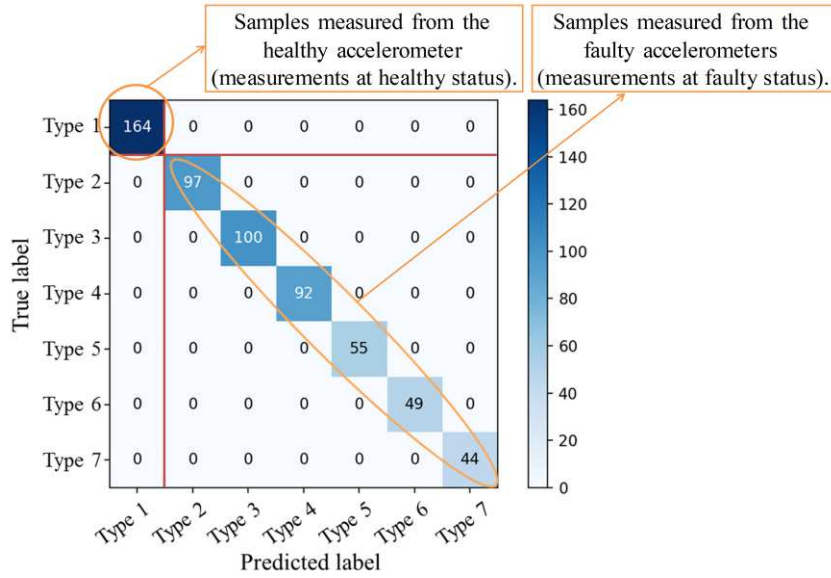| Indicators	Fault detection	Fault diagnosis
A (%)	100	100
F_1 (%)	100	100

Fig. 17. Confusion matrix of the fault detection and fault diagnosis results.

On the other hand, twelve typical examples selected from the recovered results are demonstrated in Fig. 18(a)-(l), in which two examples are selected for each faulty accelerometer. In order to quantify the recovery performance of the examples in Fig. 18, the R^2 values of these examples are summarized in Table 4. In Table 4, R_{T1}^2 represents the determination coefficient between the time-domain waveforms of the recovered shock signal and the reference shock signal, while R_{T2}^2 represents the determination coefficient between the time-domain waveforms of the faulty shock signal and the reference shock signal; R_{S1}^2 denotes the determination coefficient between the SRS curves of the recovered shock signal and the reference shock signal, while R_{S2}^2 denotes the determination coefficient between the SRS curves of the faulty shock signal and the reference shock signal. Besides, to further illustrate the recovery performance of our proposed methodology, the distributions of the recovery results of the testing set are demonstrated in Fig. 19.
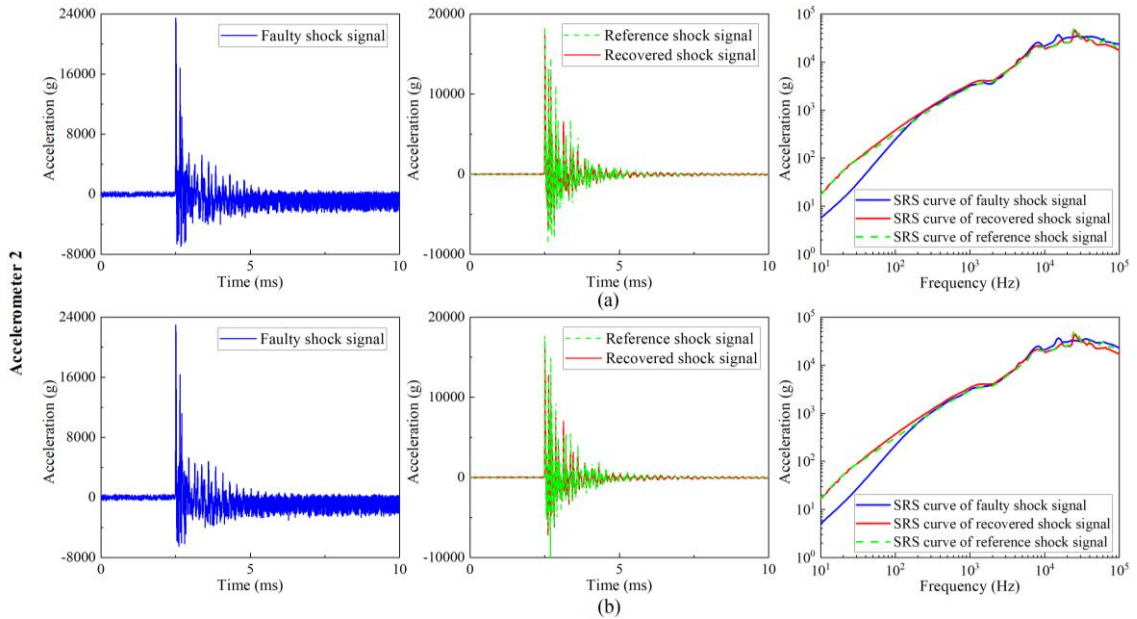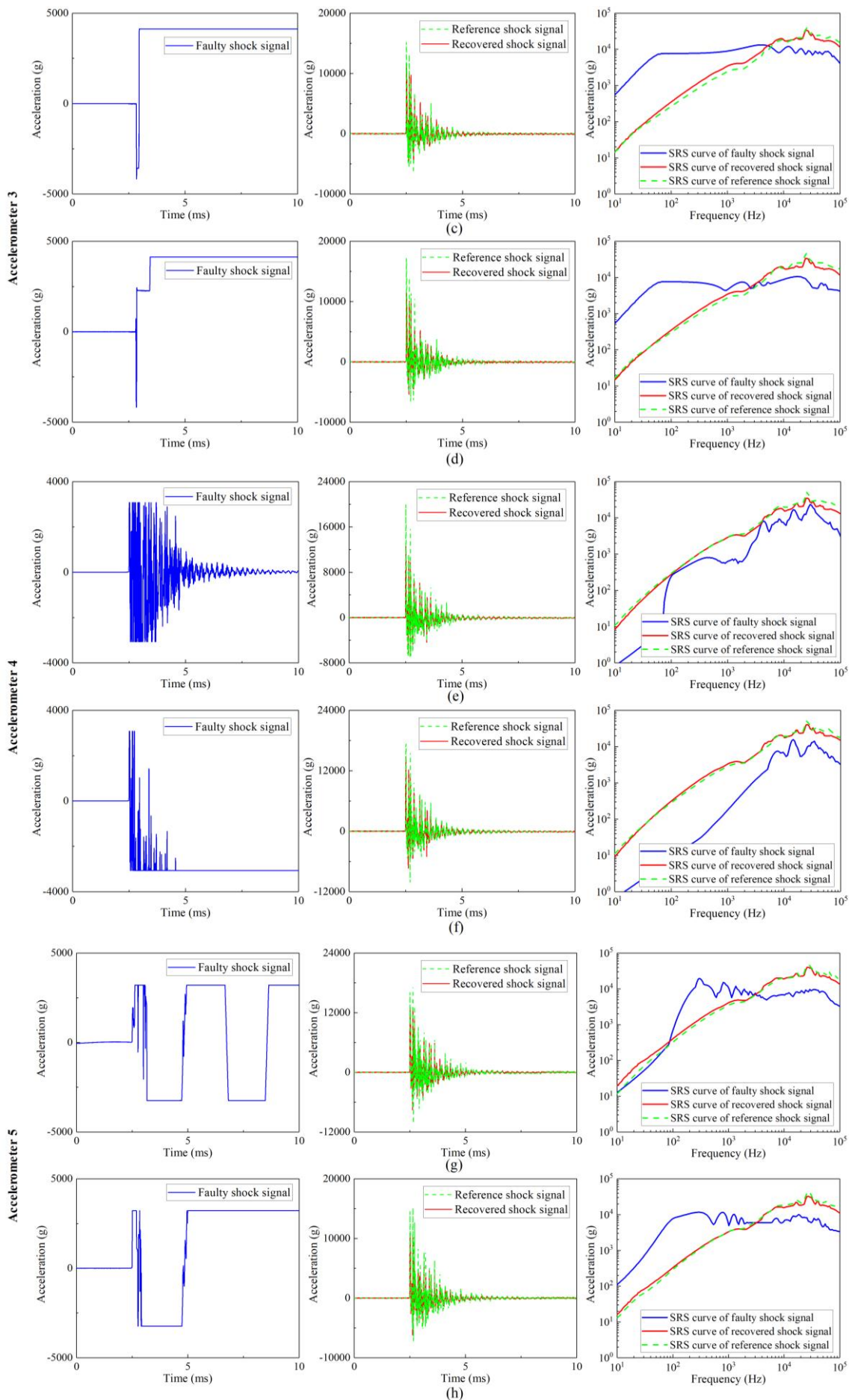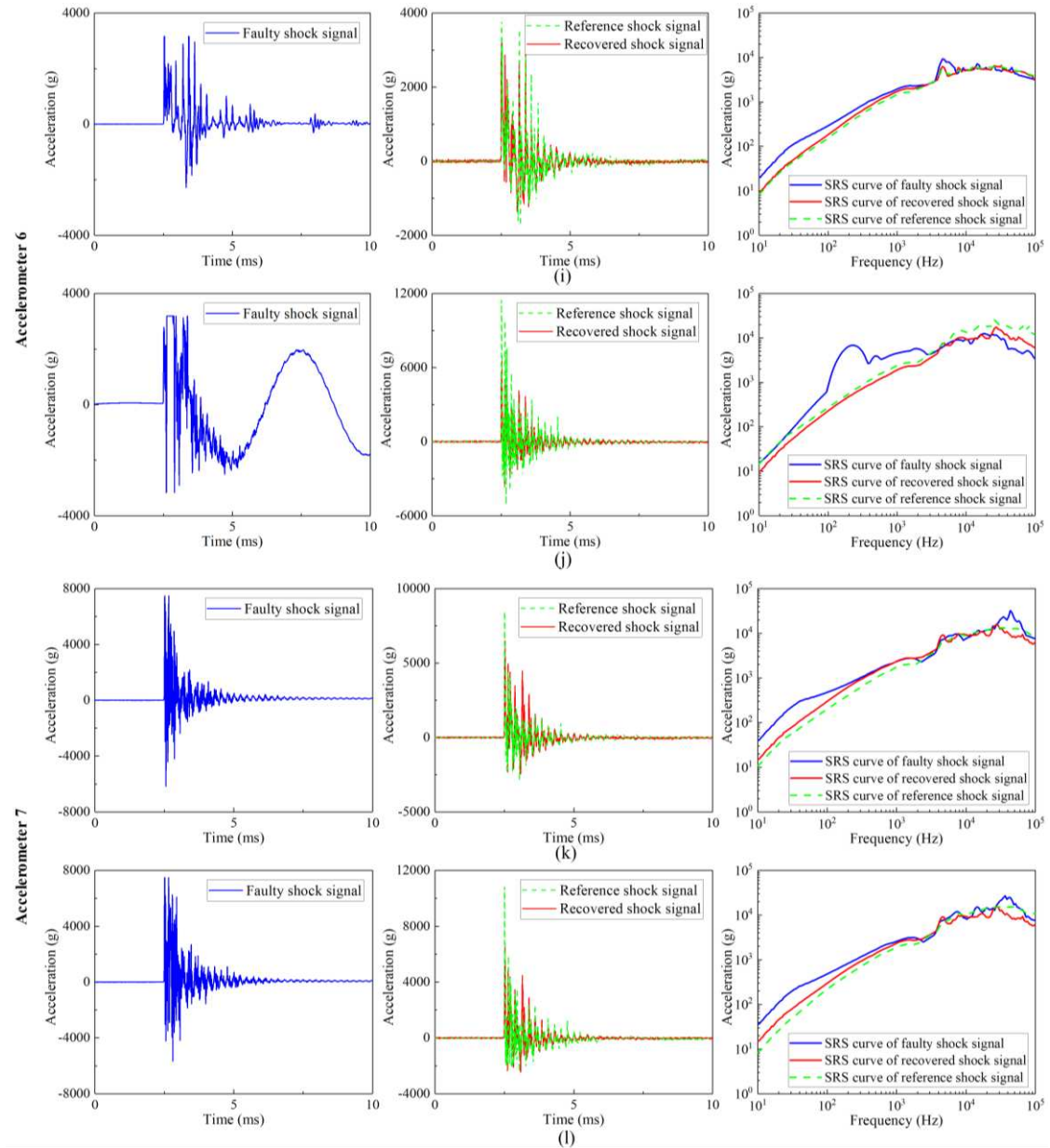
Fig. 18. Visualization of the shock signals before and after recovery. Better to be viewed in color.

Table 4. The R^2 values of the examples illustrated in Fig. 18.

Accelerometer label	Fault type	Sequence	R^2_{T1}	R^2_{T2}	R^2_{S1}	R^2_{S2}
Accelerometer 2	Type 2	Fig. 8 (a)	0.823	0.179	0.978	0.938
		Fig. 8 (b)	0.782	0.179	0.966	0.932
Accelerometer 3	Type 3	Fig. 8 (c)	0.655	-0.128	0.956	-5.597
		Fig. 8 (d)	0.645	-0.120	0.924	-7.790
Accelerometer 4	Type 4	Fig. 8 (e)	0.778	-0.857	0.899	0.159
		Fig. 8 (f)	0.764	-0.116	0.947	-0.631
Accelerometer 5	Type 5	Fig. 8 (g)	0.844	-0.129	0.968	-2.730
		Fig. 8 (h)	0.835	-0.090	0.951	-4.896
Accelerometer 6	Type 6	Fig. 8 (i)	0.657	-0.034	0.987	0.878
		Fig. 8 (j)	0.545	0.006	0.764	0.410
Accelerometer 7	Type 7	Fig. 8 (k)	0.400	0.038	0.911	0.667
		Fig. 8 (l)	0.483	0.063	0.847	0.718

Accelerometer 2
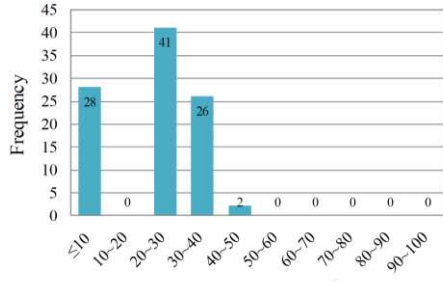
(a) The distribution of the R_{T2}^2 (%)
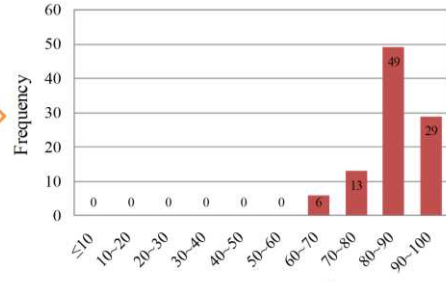
(b) The distribution of the R_{T1}^2 (%)
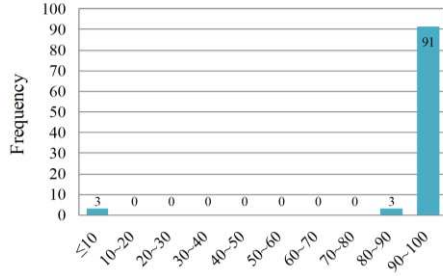
(c) The distribution of the R_{S2}^2 (%)
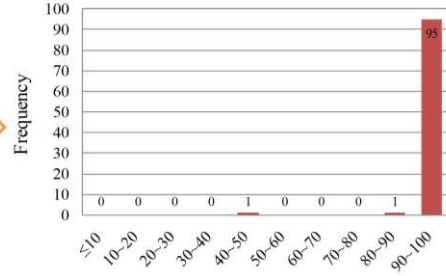
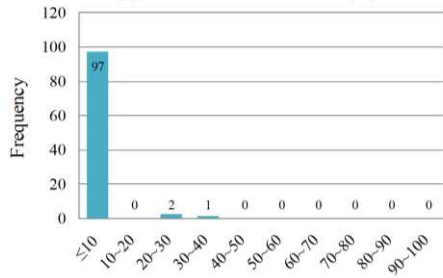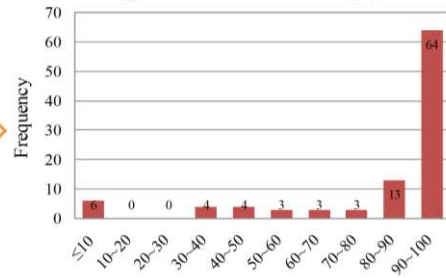(d) The distribution of the R_{S1}^2 (%)
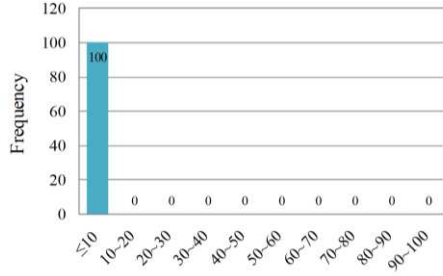
Accelerometer 3

(e) The distribution of the R_{T2}^2 (%)
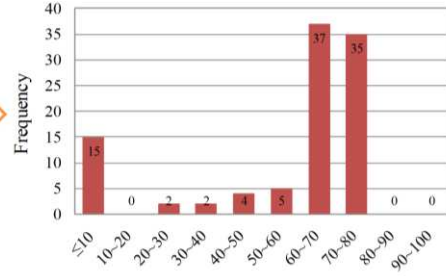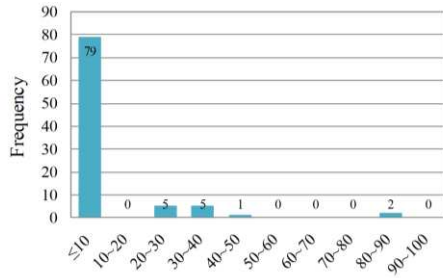
(f) The distribution of the R_{T1}^2 (%)
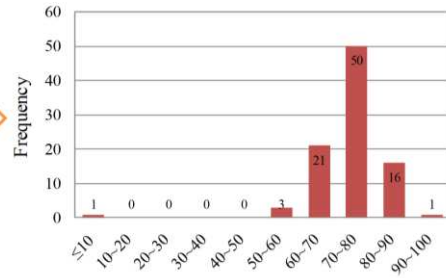
(g) The distribution of the R_{S2}^2 (%)
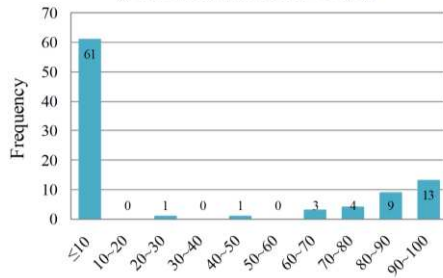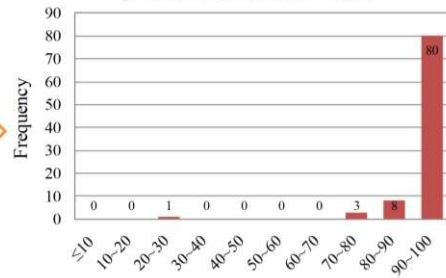
(h) The distribution of the R_{S1}^2 (%)

Accelerometer 4

(i) The distribution of the R_{T2}^2 (%)
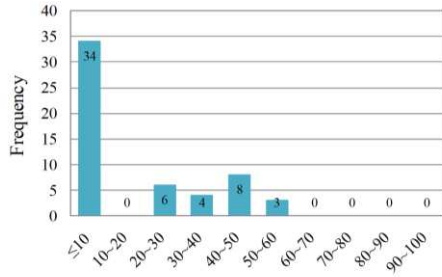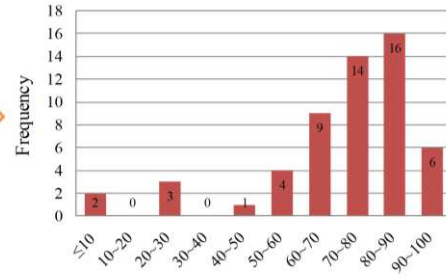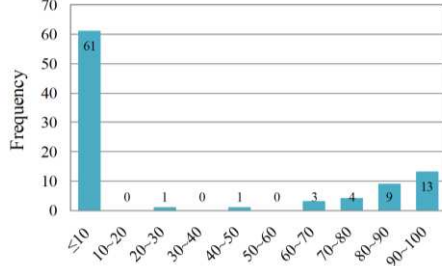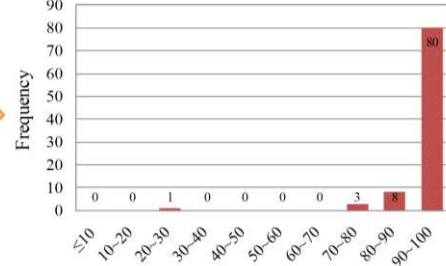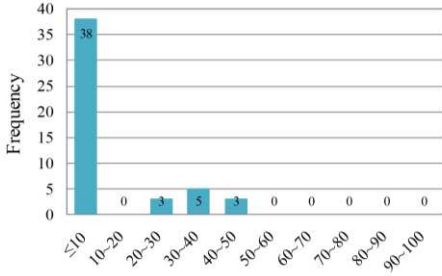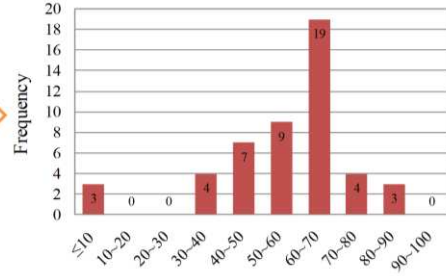
(j) The distribution of the R_{T1}^2 (%)

(k) The distribution of the R_{S2}^2 (%)
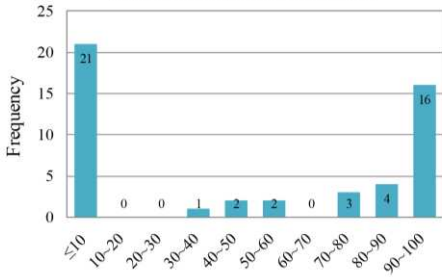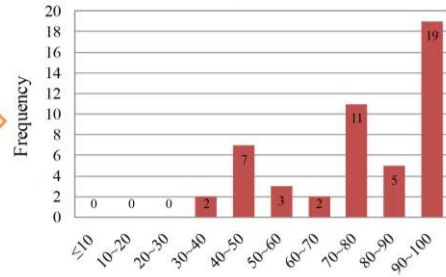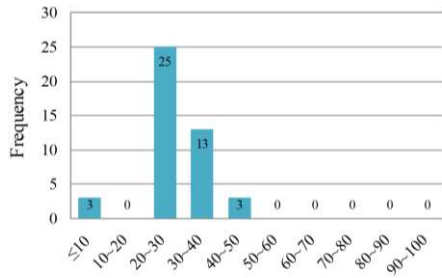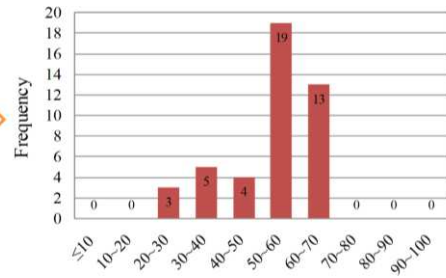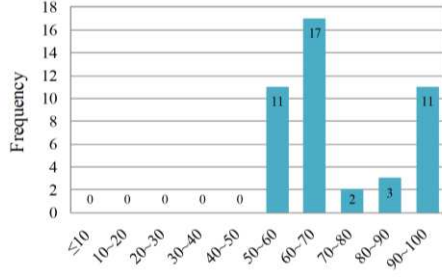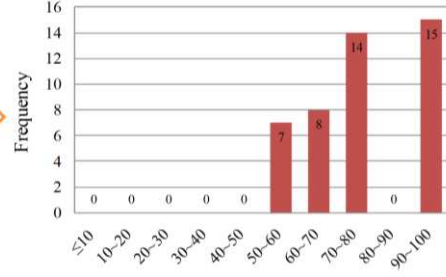
(l) The distribution of the R_{S1}^2 (%)

Fig. 19. The distributions of the performance metrics of the recovered results.

Lastly, according to Eq. (12)-(14), validated uncertainties of the recovery shock signals at different fault types can be calculated and demonstrated in Table 5.

Table 5. The validated uncertainties at different fault types.

δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7
0	29.886	44.435	34.489	41.147	43.901	46.267

5. Discussion

It can be found from Fig. 17 that all the samples fall on the diagonal blocks of the confusion matrix, which demonstrates the good performance of the proposed fault detection and the fault diagnosis methods. Besides, as shown in Table 3, the fault detection and fault diagnosis performance of the proposed ensemble learning model are quantified with two indicators: accuracy and F_1-score. It can be seen that the proposed fault detection and fault diagnosis methods both achieve good results with both the two indicators: all reaching 100%, demonstrating that the proposed deep learning-based strategy can be used to solve the fault detection and the fault diagnosis problems of high-g accelerometers effectively.

As shown in Fig. 18, the faulty accelerometers output corrupted high-g shock signals, and, meanwhile, the SRS curves of the faulty shock signals are clearly deviating from the SRS curve of the reference shock signal. By comparing the recovered shock signals and the reference shock signals, it can be observed that the shock signals recovered by the proposed DSAE-based strategy match very well with their corresponding reference shock signals. On the other hand, the SRS curves of the recovered shock signals and the SRS curves of the corresponding reference shock signals also match closely. This not only indicates that the faulty shock signals can be effectively recovered by the proposed method, but also verifies that the recovered signals can be applied into the pyroshock test directly and accurately.

The R^2 values of the results in Fig. 18 are summarized in Table 4. It can be seen from Table 4 that most of the R_{T1}^2 values are above 60%, with few examples ranging from 40% to 54.5%. Considering the high data-loss of the faulty shock signals (most of the R_{T2}^2 values are below 20%), it can be considered that the accuracy of the recovered signals in the time domain is averagely satisfactory. Additionally, it can also be found from Table 4 that all the R_{S1}^2 values are above 76.4%, and most of the R_{S1}^2 values are above 90%, which is very close to 1, showing a highly fitting accuracy between the SRS curves of the recovered signals and the SRS curves of the reference signals. Comparing the R_{S1}^2 values and the R_{S2}^2 values, it can be seen that the proposed deep learning-based method has high performance in recovering the SRS curves of the faulty shock signals.

To further illustrate the effectiveness of the proposed model, the distributions of the recovering results of the test dataset for all the faulty high-g accelerometers are demonstrated in Fig. 19. It can be seen from Fig. 19 that most of the R_{T1}^2 values are higher than 60%, while most of the R_{T2}^2 values are less than 20%, which implies that the proposed network is capable of predicting the overall time-domain waveforms of the shock signals to some extent. Similarly, as demonstrated in Fig. 19, most of the R_{S1}^2 values are above 80%, which also confirms the validity of applying our proposed method for the pyroshock tests.

From Table 5, it can be found that except δ_1, the remaining uncertainties are all above 0, which means that except accelerometer 1, accelerometer 2-7 all present faulty statuses. δ_2 is the lowest among all the uncertainties except δ_1, such that the recovery results under fault type 2 is the most trustworthy. Likewise, it can be found from Fig. 19 that, the recovery results for accelerometer 2 is the most satisfactory. Again, except δ_1, δ_4 is the second lowest among all the uncertainties, and what can also be found from Fig. 19 is that the recovery results for fault type 4 is satisfactory. Besides, δ_3, δ_5, δ_6, δ_7 are higher than δ_2 and δ_4, so that the recovery results under fault type 3, 5, 6, and 7 are with lower credibility w.r.t the recovery results under fault 2 and 4 conditions.

6. Conclusion

In this work, a data-driven strategy is proposed to self-validate high-g accelerometers used in pyroshock tests of electronic products. The proposed data-driven-based strategy mainly combines ensemble kNNs and a DSAE-based DNN. Respectively, the ensembled kNNs are designed for fault detection and fault identification; the designed DSAE is applied to recover the faulty shock signals; and a loss metric is designed to evaluate the uncertainties of the recovery shock signals under different fault types.

Lastly, a dataset comprises 3004 sets of shock signals, which are produced by seven high-g accelerometers (with six faulty accelerometers and a healthy one), are collected as the dataset, and the proposed self-validation strategy is benchmarked on this dataset. The experimental results demonstrate the validity of our proposed data-driven method, which can provide an effective self-validation solution for high-g accelerometers used in pyroshock tests. Namely, with this self-validating strategy, five output parameters based on self-validating sensor definition, i.e. original measurements, validated measurements, validated uncertainty, the status of measurements, and the fault types, can be generated successfully.

This work is the first attempt to introduce the self-validating sensor technique into the fields of high-g accelerometers, and also paves a new avenue for the study of self-validation sensors as well as the health management for shock test systems. Additionally, what should be noted is that the proposed data recovery method in this paper is applied to these six fault types of accelerometers, but not limited to these fault types.

7. Future research

As one of the first attempts in this area, it is inevitable that more research will be needed to further improve our work in future. Four main directions are discussed in this section:

(1) In the future work, one of the valuable points worth investigation will be the study of further improving recovery capacity in the local characteristics of shock signals, especially the recovery on the peak values and pulse widths of the faulty shock signals. Such that, the proposed self-validating strategy can recover not only the SRS curves but also the peak values and pulse widths of shock signals, and, hence, would be suitable for both the pyroshock test and the board-level drop test simultaneously.

(2) Another valuable point will be transferring the self-validation knowledge presented in this work into more application scenarios involving high-g measuring conditions. Therefore, the deep transfer learning would be focused in the future studies.

(3) Additionally, it can be found from Fig. 19 that the data recovery performance on different fault types is different, so the next valuable focus should be put on better understanding of the boundary of the DNN-based data reconstruction in terms of fault types or data corruption levels. What should be noted is that this direction would involve the interpretability of the DNN methodology, which is one of the bottlenecks in the current deep learning theory.

(4) Lastly, physics-informed learning is worthy for inputting focus, by which the underlying physics of the shock dynamics in the measuring process of high-g accelerometers can be integrated into the proposed machine learning models, and many benefits, such as reducing training costs and enhancing self-validating performance, can be expected with this methodology.

CRediT authorship contribution statement

Jingjing Wen: Conceptualization, Funding acquisition, Investigation, Methodology, Visualization, Writing-original draft. **Houpu Yao:** Formal analysis, Writing-review & editing. **Ze Ji:** Software, Writing-review &

editing. **Bin Wu:** Resources, Supervision, Project administration. **Feng Xu:** Resources, Supervision.

Declaration of competing interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] Monti, R., & Gasbarri, P. (2017). Dynamic load synthesis for shock numerical simulation in space structure design. *Acta Astronautica, 137*, 222-231.

[2] Yan, Y., & Li, Q. M. (2019). Low-pass-filter-based shock response spectrum and the evaluation method of transmissibility between equipment and sensitive components interfaces. *Mechanical Systems and Signal Processing, 117*, 97-115.

[3] Wen, J., Liu, C., Yao, H., & Wu, B. (2018). A nonlinear dynamic model and parameters identification method for predicting the shock pulse of rubber waveform generator. *International Journal of Impact Engineering, 120*, 1-15.

[4] Lee, J., Hwang, D. H., & Han, J. H. (2018). Study on pyroshock propagation through plates with joints and washers. *Aerospace Science and Technology, 79*, 441-458.

[5] Liu, F., Gao, S., Niu, S., Zhang, Y., Guan, Y., Gao, C., & Li, P. (2018). Optimal design of high-g MEMS piezoresistive accelerometer based on Timoshenko beam theory. *Microsystem Technologies, 24*(2), 855-867.

[6] Bateman, V., & Merritt, R. (2012). Validation of pyroshock data. *Journal of the IEST, 55*(1), 40-56.

[7] Edwards, T. S. (2003). An improved wavelet correction for zero shifted accelerometer data. *Shock and Vibration, 10*(3), 159-167.

[8] Wen, J., Yao, H., Wu, B., Ren, Y., & Ji, Z. (2019). A deep learning approach to recover high-g shock signals from the faulty accelerometer. *IEEE Sensors Journal, 20*(4), 1761-1769.

[9] Yao, H., Wen, J., Ren, Y., Wu, B., & Ji, Z. (2019, May). Low-cost measurement of industrial shock signals via deep learning calibration. In *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2892-2896). IEEE.

[10] Yuan, H. J., & Jiang, T. M. (2006). Analysis and treatment of measured pyrotechnic shock data. *Journal of Solid Rocket Technology, 29*(1), 72-74.

[11] M. P. Henry & D. W. Clarke. (1993). The self-validating sensor: rationale, definitions and examples. *Control Engineering Practice, 1*(4), 585-610.

[12] Feng Z., Wang Q., & Shida K. (2007). A review of self-validating sensor technology. *Sensor Review, 27*(1), 48-56.

[13] Majumder, B. D., Roy, J. K., & Padhee, S. (2018). Recent advances in multifunctional sensing technology on a perspective of multi-sensor system: a review. *IEEE Sensors Journal, 19*(4), 1204-1214.

[14] Yang, J., Lin, L., Sun, Z., Chen, Y., & Jiang, S. (2017). Data validation of multifunctional sensors using independent and related variables. *Sensors and Actuators A: Physical, 263*, 76-90.

[15] Yang, J., Chen, Y., & Zhang, L. (2017). An efficient approach for fault detection, isolation, and data recovery of self-validating multifunctional sensors. *IEEE Transactions on Instrumentation and Measurement, 66*(3), 543-558.

[16] Michael A. C., Maurice G. C., Mihaela D. et al. (2008). The application of self-validation to wireless sensor networks. *Measurement Science and Technology, 19*(12): 125201.

[17] Liu, B., Han, T., & Zhang, C. (2015). Error correction method for passive and wireless resonant SAW temperature sensor. *IEEE Sensors Journal, 15*(6), 3608-3614.

[18] Feng Z., Wang Q., & Shida K. (2007). Design and implementation of a self-validating pressure sensor. *IEEE Sensors Journal, 9*(3), 207-218.

[19] Jiang, Y., Yin, S., Dong, J., & Kaynak, O. (2020). A Review on Soft Sensors for Monitoring, Control and Optimization of Industrial Processes. *IEEE Sensors Journal*. (DOI: 10.1109/JSEN.2020.3033153, in press).

[20] Jiang, Y., & Yin, S. (2018). Recent advances in key-performance-indicator oriented prognosis and diagnosis with a MATLAB toolbox: DB-KIT. *IEEE Transactions on Industrial Informatics, 15*(5), 2849-2858.

[21] Wang, P., Yan, M., Zhang, L., & Zhang, M. (2019). A correction method for the underwater shock signals of floating shock platforms based on a combination of FFT and low-frequency oscillator. *Shock and Vibration, 2019*, 1-8.

[22] Cao H, Cui R, Liu W, et al. (2021). Dual mass MEMS gyroscope temperature drift compensation Based on TFPF-MEA-BP algorithm. *Sensor Review*. (DOI: 10.1108/SR-09-2020-0205, in press).

[23] Lee, J. R., Chia, C. C., & Kong, C. W. (2012). Review of pyroshock wave measurement and simulation for space systems. *Measurement, 45*(4), 631-642.

[24] Li, X., Zhang, W., & Ding, Q. (2019). Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks. *IEEE Transactions on Industrial Electronics, 66*(7), 5525-5534.

[25] You, G. W., Park, S., & Oh, D. (2017). Diagnosis of electric vehicle batteries using recurrent neural networks. *IEEE Transactions on Industrial Electronics, 64*(6), 4885-4893.

[26] Rabiei, E., Droguett, E. L., & Modarres, M. (2016). A prognostics approach based on the evolution of damage precursors using dynamic Bayesian networks. *Advances in Mechanical Engineering, 8*(9), 1-19.

[27] Wen, J., Yao, H., Ji, Z., Wu, B., & Xia, M. (2021). On fault diagnosis for high-g accelerometers via data-driven models. *IEEE Sensors Journal, 21*(2), 1359-1368.

[28] Gao, Z., Cecati, C., & Ding, S. X. (2015). A survey of fault diagnosis and fault-tolerant techniques—Part II: Fault diagnosis with knowledge-based and hybrid/active-based approaches. *IEEE Transactions on Industrial Electronics, 62*(6), 3768-3774.

[29] Yu, Y., Yao, H., & Liu, Y. (2019). Aircraft dynamics simulation using a novel physics-based learning method. *Aerospace Science and Technology, 87*, 254-264.

[30] Neerukatti, R. K., Chattopadhyay, A., Iyyer, N., & Phan, N. (2018). A hybrid prognosis model for predicting fatigue crack propagation under biaxial in-phase and out-of-phase loading. *Structural Health Monitoring, 17*(4), 888-901.

[31] Jahani, S., Setarehdan, S. K., Boas, D. A., & Yücel, M. A. (2018). Motion artifact detection and correction in functional near-infrared spectroscopy: a new hybrid method based on spline interpolation method and Savitzky–Golay filtering. *Neurophotonics, 5*(1), 015003.

[32] Wu, B., Liu, C., & Wen, J. (2017, October). The optimized algorithm for working parameters of the vertical impact testing machine. In *2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)* (pp. 424-430). IEEE.

[33] Meng, J., & Dasgupta, A. (2017). Influence of secondary impact on printed wiring assemblies—Part II: Competing failure modes in surface mount components. *Journal of Electronic Packaging, 139*(3), 031001.

[34] Zou, H., Wang, J., Chen, F., Bao, H., Jiao, D., Zhang, K., ... & Li, X. (2017). Monolithically integrated tri-axis shock accelerometers with MHz-level high resonant-frequency. *Journal of Micromechanics and Microengineering, 27*(7), 075009.

[35] Narasimhan, V., Li, H., & Jianmin, M. (2015). Micromachined high-g accelerometers: a review. *Journal of Micromechanics and Microengineering, 25*(3), 033001.

[36] Mishra, R. K. (2018). A review on fracture mechanics in piezoelectric structures. *Materials Today: Proceedings, 5*(2), 5407-5413.

[37] Walraven, J. A. (2005). Failure analysis issues in microelectromechanical systems (MEMS). *Microelectronics Reliability, 45*(9-11), 1750-1757.

[38] Cao, H., Zhang, Z., Zheng, Y., et al. (2021). A new joint denoising algorithm for high-g calibration of MEMS accelerometer based on VMD-PE-Wavelet threshold. *Shock and Vibration, 2021*(8855878), 1-16.

[39] Liu, J., Shi, Y., Li, P., Tang, J., Zhao, R., & Zhang, H. (2012). Experimental study on the package of high-g accelerometer. *Sensors and Actuators A: Physical, 173*(1), 1-8.

[40] Wu, M. L., & Lan, J. S. (2018). Reliability and failure analysis of SAC 105 and SAC 1205N lead-free solder alloys during drop test events. *Microelectronics Reliability, 80*, 213-222.

[41] García-Pérez, A., Sorribes-Palmer, F., Alonso, G., & Ravanbakhsh, A. (2019). FEM simulation of space instruments subjected to shock tests by mechanical impact. *International Journal of Impact Engineering, 126*, 11-26.

[42] Zhao, H., Ding, J., Zhu, W., Sun, Y., & Liu, Y. (2019). Shock response prediction of the typical structure in spacecraft based on the hybrid modeling techniques. *Aerospace Science and Technology, 89*, 460-467.

[43] Irvine, T. (2002). An introduction to the shock response spectrum. *Rev P, Vibrationdata*.

[44] Morais, O. M. F., & Vasques, C. M. A. (2017). Shock environment design for space equipment testing. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 231*(6), 1154-1167.

[45] Meng, J., Douglas, S. T., & Dasgupta, A. (2016). MEMS packaging reliability in board-level drop tests under severe shock and impact loading conditions–part I: Experiment. *IEEE Transactions on Components, Packaging and Manufacturing Technology, 6*(11), 1595-1603.

[46] Zhang, Y., Cao, G., Wang, B., & Li, X. (2019). A novel ensemble method for *k*-nearest neighbor. *Pattern Recognition, 85*, 13-25.

[47] Yuan, X., Huang, B., Wang, Y., Yang, C., & Gui, W. (2018). Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE. *IEEE Transactions on Industrial Informatics, 14*(7), 3235-3243.

[48] Jelinek F, Bahl L, Mercer R. (1975). Design of a linguistic statistical decoder for the recongnition of continuous speech. *IEEE Transactions on Information Theory, 21*(3): 250-256.

[49] Zabalza, J., Ren, J., Zheng, J., Zhao, H., Qing, C., Yang, Z., ... & Marshall, S. (2016). Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing, 185*, 1-10.

[50] Jiao, J., Zhao, M., Lin, J., & Ding, C. (2019). Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis. *IEEE Transactions on Industrial Electronics, 66*(12), 9858-9867.

[51] Jiao G., He H., Yan J., et al. (2019). Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox. *IEEE Transactions on Industrial Electronics, 66*(4), 3196-3207.

[52] Ferraro, M. B., Colubi, A., Gonzalez-Rodriguez, G., & Coppi, R. (2011). A determination coefficient for a linear regression model with imprecise response. *Environmetrics, 22*(4), 516-529.

[53] Shao, H., Jiang, H., Zhao, H., & Wang, F. (2017). A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing, 95*, 187-204.

[54] Lei, Y., Jia, F., Lin, J., Xing, S., & Ding, S. X. (2016). An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Transactions on Industrial Electronics, 63*(5), 3137-3147.

[55] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing, 95*, 213-237.