

PAPER

Detection of impact on aircraft composite structure using machine learning techniques

To cite this article: Li Ai *et al* 2021 *Meas. Sci. Technol.* **32** 084013

View the [article online](#) for updates and enhancements.

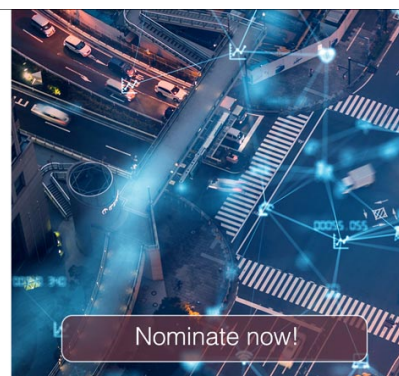


The Electrochemical Society
Advancing solid state & electrochemical science & technology

The ECS is seeking candidates to serve as the
Founding Editor-in-Chief (EIC) of ECS Sensors Plus,
a journal in the process of being launched in 2021

The goal of ECS Sensors Plus, as a one-stop shop journal for sensors, is to advance the fundamental science and understanding of sensors and detection technologies for efficient monitoring and control of industrial processes and the environment, and improving quality of life and human health.

Nomination submission begins: May 18, 2021



Detection of impact on aircraft composite structure using machine learning techniques

Li Ai¹ , Vafa Soltangharai¹, Mahmoud Bayat¹ , Michel Van Tooren^{2,*} and Paul Ziehl¹

¹ Department of Civil and Environmental Engineering, University of South Carolina, Columbia, SC, United States of America

² Collins Aerostructures, Chula Vista, CA, United States of America

E-mail: ziehl@cec.sc.edu

Received 28 September 2020, revised 4 February 2021

Accepted for publication 18 February 2021

Published 19 May 2021



Abstract

Aircraft structures are exposed to impact damage caused by debris and hail during their service life. One of the design concerns in composite structures is the resistance of layered surfaces to damage, which occurs from impacts with various foreign objects. Therefore, the impact localization and damage quantification of impacts should be studied and considered to address flight safety and to reduce costs associated with a regularly scheduled visual inspection. Since the structural components of the aircraft are large scale, visual inspection and monitoring are challenging and subject to human error. This paper presents a promising solution that can automatically detect and localize an impact that may occur during flight. To achieve this goal, acoustic emission (AE) is employed as an impact monitoring approach. Random forest and deep learning were adopted for training the source location models. An AE dataset was collected by conducting an impact experiment on a full-size thermoplastic aircraft elevator in a laboratory environment. A dataset consisting of AE parametric features and a dataset consisting of AE waveforms were assigned to a random forest classifier and deep learning network for the investigation of their applicability of impact source localization. The results obtained were compared using the source localization approach in previous research using a conventional artificial neural network. The analysis of results shows the random forest and deep learning leads to better event localization performance. In addition, the random forest model can provide the importance of features. By deleting the least important features, the storage required to save the input and the computing time for the random forest is greatly reduced, and an acceptable localization performance can still be obtained.

Keywords: impact monitoring, acoustic emission, artificial neural network, random forest, stacked autoencoder

(Some figures may appear in colour only in the online journal)

1. Introduction

Impact events are inevitable in the operation of aircraft. It is necessary to conduct prompt maintenance in order to ensure

safety operability. Conventional inspections of aircraft structures usually refer to visual inspections, which are conducted on the apron during the gaps between flights, and more detailed inspections are performed at the base, where more sophisticated methods can be applied [1]. A nondestructive health monitoring system can be used to measure the impact location instead of or in addition to workers to assist the visual

* This paper was conducted when Michel van Tooren was working at the University of South Carolina

inspections and to mark the damaged zones for further detailed measurement of structural health condition. Data collected by such systems can be downloaded and processed between flight operations. Once proven effective, the nondestructive health monitoring system may improve the accuracy of inspection, shorten the time for inspection, and extend the time on maintenance and repair.

Acoustic emission (AE) is a non-destructive structural health monitoring and sensing technology [2–11]. This method is sensitive and has continuous monitoring capabilities [12–14]. It has been utilized as a measurement tool to assess the health condition of materials [15–18]. Santos-Leal *et al* [15] developed a simultaneous measurement system to investigate the relation between the information measured by AE and electrical resistance variation during the stress corrosion crack propagation in high-strength low-alloy steels. The results indicated the cumulative counts and cumulative energy of AE showed a very similar temporal behavior with electrical resistance variation. Li *et al* [16] proposed an approach to quantify the fatigue cracking size in rail tracks using AE. An empirical model derived was utilized for the quantification of crack size using the AE wave generated by the crack closure process during the fatigue test. The results indicated that the AE count rate has correlations with the crack length. Liu *et al* and Du *et al* [17] utilized AE as a measurement method towards the forecasting and prevention of disastrous mining accidents such as rock bursts. The particle flow code (PFC2D) is developed by the authors to study the hidden patterns among the damage mechanism of the rock and AE signals. Liu *et al* [18] investigated AE signals generated by coal ruptures under uniaxial compressive loads. The relationship between the coal damage and the AE counts, frequency, and AE energy was studied. Based on the literature, AE can be used as a measurement technique in structural health monitoring in different applications.

Developing AE monitoring systems compatible with the requirements of aircraft is of interest. The traditional localization approach is the time of arrival method. The AE source localization can be obtained by analyzing the signals captured by multiple AE sensors. However, due to weight and power restrictions, the main challenge of applying AE in aircraft is to use as few sensors as possible while obtaining accurate impact localization results. Machine learning techniques can be a tool to solve the source localization problem in this situation. The authors proposed a passive structural health monitoring system for the measurement of impact location on aircraft elevators in a previous paper [19]. AE signals caused by impacts are collected by a single AE sensor. Parametric features such as count number, signal strength, and duration were employed as the input data set of the backpropagation (BP) artificial neural network (ANN). The outputs of the neural network were zonal source localization results. The results showed that the proposed system is feasible and is able to obtain an acceptable accuracy in measuring the location of impact. Compared to the traditional source location approach, a machine learning-based method like ANN can receive good localization accuracy by using only one AE sensor, which is suitable for an aircraft. However, a major problem of ANN in this

localization approach is the convergence, as well as becoming easily trapped in locally optimal solutions [20]. When the number of layers in a neural network exceeds four, the optimization of the entire network would be a problem because of the vanishing/explosion gradient [21]. In addition, another major problem is that it is necessary to manually extract features on the collected data and select appropriate features as training input [22]. This usually depends on experience and other important features may be overlooked. An improved passive health monitoring system using AE with an advanced source localization approach is needed.

Random forest is an approach that can be used to improve passive health monitoring systems. The random forest is a machine learning method based on statistical learning theory [23]. It combines the bootstrap resampling method and a decision tree algorithm. The essence of this algorithm is to train multiple decision tree models independently, then put together the results of these decision trees. The result with the most votes is the final prediction result. The random forest has an advantage in that the importance of input variables can be ranked. It provides theoretical support for feature selection. In recent years, the random forest has been successfully applied as an analysis approach tool in AE monitoring. Shevchik *et al* [24] successfully developed an approach to predict the scuffing failure in lubricated mechanical components by using AE and random forest. Wang *et al* [25] employed the random forest machine learning approach to study the link between AE signals and the phenomena of fiber fracture during natural fiber reinforced plastic machining. Iquebal *et al* [26] utilized random forest to connect the patterns of the corresponding AE signals with the microstructural phases on a metallic workpiece surface under a nanoindentation-based lithography process. These studies show the capability of random forest in the field of AE monitoring.

An alternative approach is adopting deep learning. Deep learning is a method based on representational learning of data in machine learning [27]. The concept of deep learning stems from the study of ANNs. Deep learning combines low-level features to form more abstract high-level features to discover the distributed feature representations of data. Deep learning was initially proposed by Hinton *et al* [28]. In this study, the authors employed the greedy layer-wise training technique and unsupervised back-propagation algorithm to solve the problem of vanishing/explosion gradient and obtained good results. Another advantage of deep learning is that the raw data can be employed as the input data which means manually feature extraction is not needed [29]. Recently, deep learning has been utilized in the fields of vibration signal processing and AE monitoring [30–33]. Li *et al* [30] proposed a gearbox fault diagnosis method based on AE. A deep random forest was utilized as the classification algorithm. He *et al* [31] used a deep learning model to classify AE signals to diagnose bearing failures. Shevchik *et al* [32] used AE technology to develop an on-site health monitoring system for additive manufacturing. An AE data classification method based on a spectral convolutional neural network was developed. Ebrahimkhanlou *et al* [33] worked on a deep learning framework for locating AE events

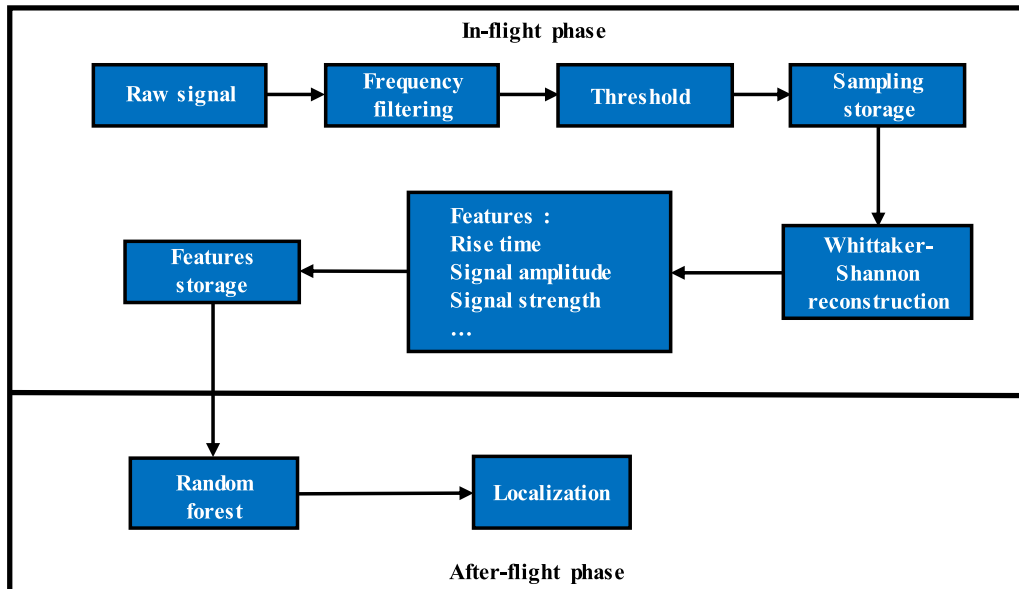


Figure 1. Procedures of the random forest-based passive monitoring system.

on the metal-like plate structures. These studies demonstrate the potential of applying deep learning in the field of AE monitoring.

The main contribution of this paper is to develop an optimized algorithm for precisely estimating impact locations using deep learning and random forest. Currently, the authors are not aware of published work where random forest and deep learning is applied to the measurement of impact locations in realistic aerospace structural components. To fill this gap, an improved AE-based passive structural health monitoring system using random forest and deep learning is developed herein and the source location results are compared with the previously published method. The proposed system can facilitate structural health monitoring of large-scale aerospace structures by improving the source localization measurement and reducing human error.

2. Theoretical background and methodology

2.1. Zonal source location using random forest algorithm

This study focuses on an improved random forest-based passive monitoring system. The random forest-based passive monitoring system includes two phases: the in-flight phase and the after-flight phase. The different steps of these two phases are identified in the flowchart which is presented in figure 1. When the aircraft is in the in-flight phase, one single AE sensor is mounted on the spar of the elevator component for the recording of impact events during the flight operation. The initial signals received by the AE sensor are analog signals, which can be analyzed in a computer platform after sample processing. Before the initial AE signals are sampled, they are filtered by a low-frequency high-pass filter which is used to avoid aliasing during processing, and a high-frequency low-pass filter which is used to match the sampling rate obtained

by the data acquisition system. Further processing of the signal is dependent upon a settable threshold for the signal amplitude (voltage). The processed signals are stored after sampling.

The stored signals are employed to localize the impact events. Before the training of random forest, features are extracted from these stored signals. The stored signals are reconstructed using the Whittaker–Shannon reconstruction concept [34]. AE features such as rise time, signal amplitude and signal strength can be extracted from the AE wave after reconstruction. The features are then gathered as a final data set for after-flight analysis.

The random forest model is implemented on the ground, after landing, and uses the stored features to localize each AE event.

2.1.1. Whittaker–Shannon reconstruction. The Whittaker–Shannon reconstruction algorithm reconstructs the original AE signals from a set of sampling data points. The calculation process is presented in equation (1):

$$X(t) = \sum_{n=-\infty}^{\infty} X_s(n) \operatorname{sinc}\left(\frac{t}{T_s} - n\right) \quad (1)$$

where X_s refers to the sampling spectrum. It can be obtained by equation (2):

$$X_s(\omega) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X\left(\frac{\omega - 2\pi k}{T_s}\right) \quad (2)$$

where T_s refers to the sampling period and $\operatorname{sinc}(\ast)$ is the inverse Fourier transform of the ideal low pass filter G , which must satisfy the following requirement in equation (3):

$$G = \operatorname{sinc}\left(\frac{t}{T_s}\right) = \begin{cases} 1, & t = 0 \\ \frac{\sin\left(\frac{\pi t}{T_s}\right)}{\pi t/T_s}, & t \neq 0 \end{cases} \quad (3)$$

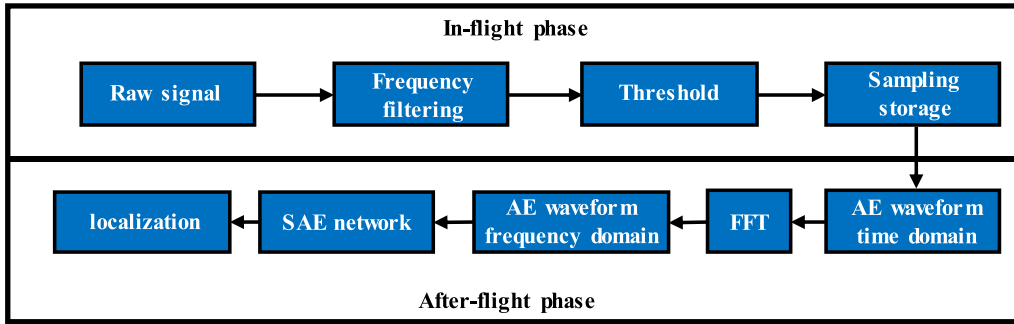


Figure 2. Procedures of DNN-based passive monitoring system.

2.1.2. *Random forest.* Random forest is an ensemble learning algorithm, a kind of bagging algorithm [23]. This algorithm provides the results by combining multiple weak learning models. The results of the weak learning models are voted or averaged to obtain the result of the overall model [35]. The weak learning model refers to the model with a prediction accuracy that is only slightly higher than the random guess. The weak learning model employed in the random forest is the decision tree.

The decision tree makes classification decisions based on multiple features. At each node of the tree, the leaf node of the next layer is branched through a criterion according to the performance of the features. With layer-by-layer branching, the sample categories included in the leaf nodes will gradually become consistent, and the terminal leaf node is the classification result of the decision trees. In this paper, the Gini impurity of the node is used as the branching criterion when generating the decision tree. The Gini impurity of a node refers to the probability that a sample randomly selected from a node is misclassified when the sample is classified according to the distribution of the samples in the node. Therefore, the purity of the samples is negatively correlated with the Gini impurity.

Assuming that the sample set N contains K categories, then the Gini impurity of node t is obtained by equation (4):

$$\text{Gini}(S) = 1 - \sum_{i=1}^K P(i/t) \tag{4}$$

where $P(i/t)$ is the probability of category i at node t . When $\text{Gini}(S) = 0$, or less than a predetermined threshold, it means that the samples belong to the same category. Otherwise, the sample is divided into two parts $N1$ and $N2$, according to feature F , and then allocated to the two sub-nodes. As shown in equation (5):

$$\text{Gini}(N, F) = \frac{N1}{N} \text{Gini}(N1) + \frac{N2}{N} \text{Gini}(N2). \tag{5}$$

According to this layer-by-layer branching, until the number of samples in the node is less than the predetermined threshold, or the Gini impurity of the sample set is less than the predetermined threshold, or there are no more features, the system stops growing and forms a decision tree to accomplish classification and prediction.

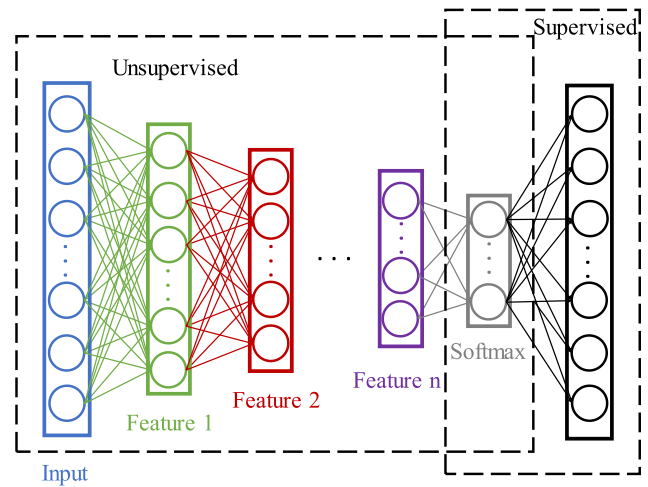


Figure 3. A stacked autoencoder contains n autoencoders.

Bagging is a parallel ensemble learning method. Based on the bootstrapping method, a fixed number of samples are collected from the training set with replacements. Thereby, a sample set for each basic learning model is formed. Because of the replacement sampling, some samples may be repeated, while some samples may not be drawn. The final result is voted or averaged from the result of all the basic models. The random forest algorithm is a combination of a decision tree and bagging. The decision tree is utilized as the basic learning model. Bagging improves the generalization error by reducing the variance of the basic learning model. The performance of bagging depends on the stability of the basic model. When the basic model is unstable, bagging helps to reduce the error caused by the random fluctuation of the training set. If the basic model is stable, bagging does not improve the performance of the model, and may even reduce the model's performance. The decision tree plus the bagging effectively decreases the variance of a single decision tree, thereby obtaining a complete random forest.

The random forest model can calculate and evaluate the importance of features through the feature division process while predicting or classifying [35]. The calculation requires the help of the Gini impurity when the leaf node is branching; the method is shown as equations (6) and (7):

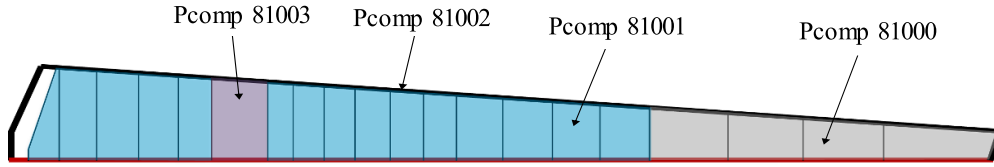


Figure 4. Elevator specimen.

Table 1. Laminate lay-up.

| Pcomp | Description | Lay-up |
|--------|---|--|
| 81 000 | 5-ply skin panel | (45, 0/90, 45, 0/90, 45) |
| 81 002 | 7-ply skin at rear beam and rib interface | (45, 0/90, 45, 0/90, 45, 0/90, 45) |
| 81 003 | 8-ply skin panel | (45, 0/90, 45, 0/90, 0/90, 45, 0/90, 45) |

$$I_t(F) = \text{Gini}(N) - \text{Gini}(N, F) \quad (6)$$

$$S(F) = \sum_t I_t(F) \quad (7)$$

where $I_t(F)$ refers to the decrease of the Gini impurity before and after node t is split into two sub-nodes according to feature F . The absolute importance of feature $S(F)$ can be defined as the sum of $I_t(F)$ at all nodes split by feature F . The importance score of each feature can be obtained by normalizing the absolute importance of all features.

2.2. Zonal source location using deep neural network

In addition to random forest, another passive monitoring system based on deep learning is investigated in this study. As in the system above, it contains the in-flight phase and the after-flight phase. The workflow of the system is shown in figure 2. The difference with the system using random forest is that there is not feature extraction. In the after-flight phase, the collected raw AE signals are input into the deep learning model for impact localization. The deep learning model utilized in this system is a stacked autoencoder (SAE).

2.2.1. SAE. An autoencoder is a typical ANN which usually contains three layers: the input layer, the hidden layer, and the output layer. The number of neurons in the input and output layers are kept consistent, while the number of neurons in the hidden layer is usually less than the input and output layers. The autoencoder algorithm compresses the input data according to the number of neurons of the hidden layer and reconstructs the output of compressed data to the output layer [36]. The compression of input data can be considered as the extraction of features. The SAE is a deep learning algorithm based on the stacking technique [37]. The SAE is composed of multiple autoencoders. The compressed features obtained

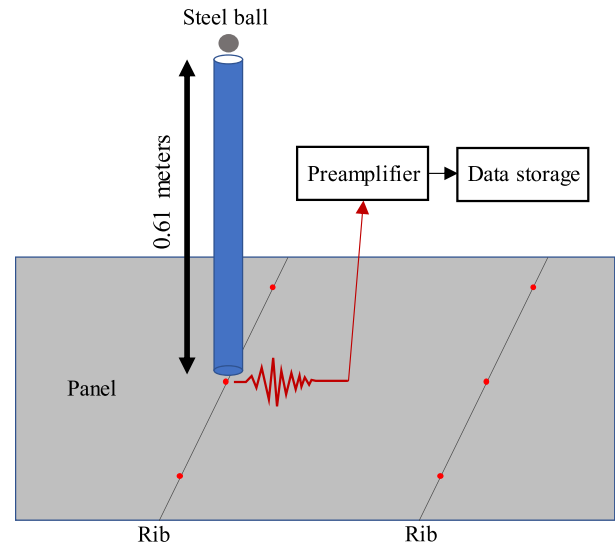


Figure 5. Steel ball dropping test.

by the previous autoencoder are employed as the input data of the following autoencoder.

Assuming the input data for the input layer is defined as an n -dimensional vector, the process of mapping the input data to the k ($k < n$) dimensional vector in the hidden layer is named as the encoder stage. The process that reconstructs the k -dimensional vector in the hidden layer to a new n -dimensional vector in the output layer is named as the decoder stage. The process of encoder and decoder stages can be expressed by equations (8) and (9):

$$G = f(w_{(i)}^e x^i + b_{(i)}^e) \quad (8)$$

$$\hat{x}^i = f(w_{(i)}^d x^i + b_{(i)}^d) \quad (9)$$

where G refers to the k -dimensional compressed feature in the hidden layer, x^i is the i th input data from the input dataset, $w_{(i)}^e$ and $b_{(i)}^e$ are the weight and bias in the encoder stage, \hat{x}^i refers to the i th output data in the output layer, $w_{(i)}^d$ and $b_{(i)}^d$ are the weight and bias in the decoder stage, and $f(*)$ is the sigmoid activation function which is utilized for transforming the set of neurons in the previous layer into a given neuron in the subsequent layer [38]. The function is presented in equation (10):

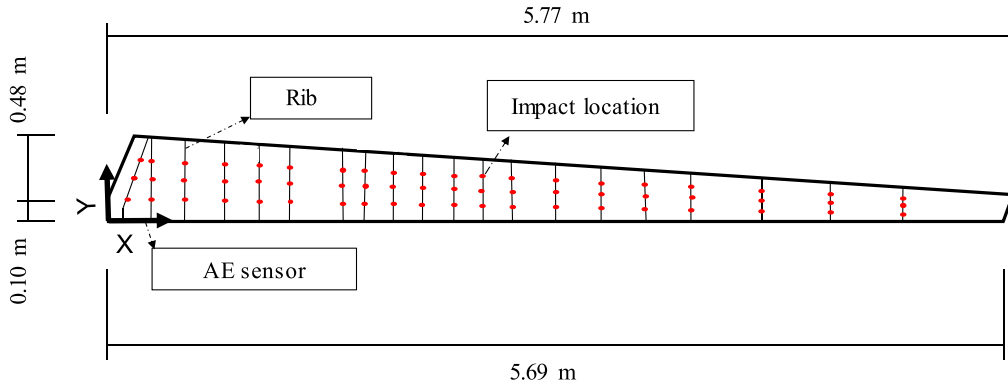


Figure 6. Impact and sensor locations.

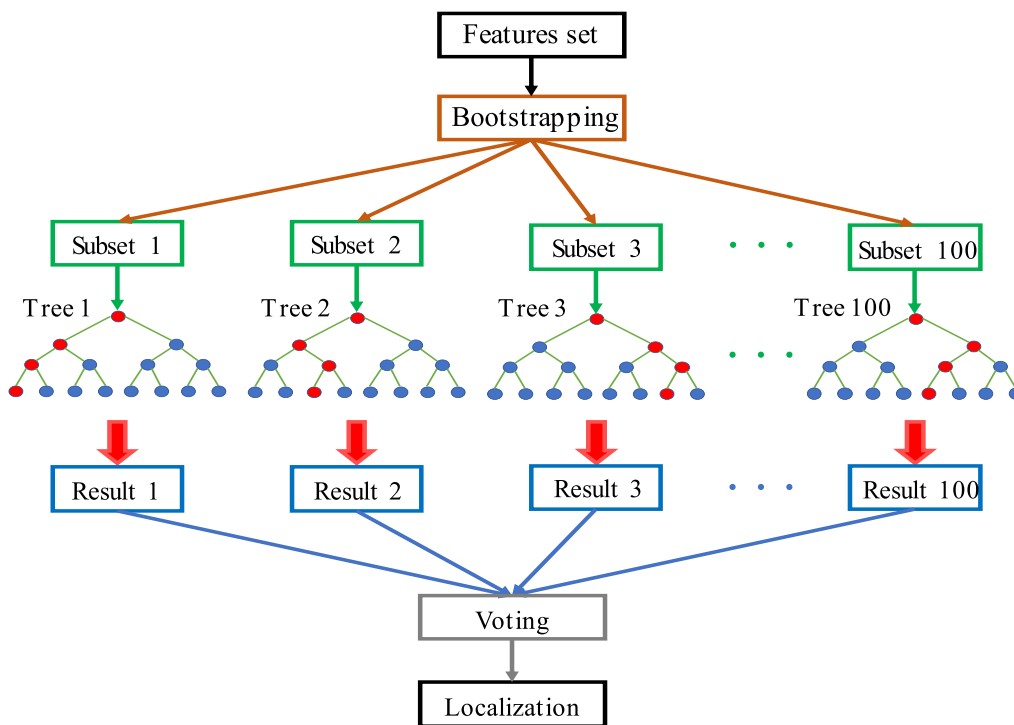


Figure 7. The architecture of the random forest.

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{10}$$

3. Experimental description

3.1. Experimental setup

The training object of the autoencoder is to minimize the error between the input data and the output data. An objective function is designed. The optimal parameter set $\{w_{(i)}^e, b_{(i)}^e, w_{(i)}^d, b_{(i)}^d\}$ will be determined in the iterations of the training process. During the training of SAEs, several autoencoders are utilized to compress the data. In other words, the data is compressed several times by multiple autoencoders. The last autoencoder is connected by a SoftMax layer which is intended to do the classification based on the final compressed features. The training process of multiple autoencoders is unsupervised, while the process of classification by the SoftMax layer is supervised by the labels. The workflow of an SAE is shown in figure 3.

To evaluate the performance of the systems proposed in this paper. An impact experiment was conducted in an aircraft elevator component in a laboratory setting. The elevator component was mounted on a steel frame structure, which was fabricated by the 5 inch mild steel channels (6.10 m long \times 0.61 m high). The hinge brackets on the elevator spar were connected to hinge points located on the frame. To simulate the flexure of the horizontal tail during the realistic flight operation, a turnbuckle was employed to create bending in the elevator component.

A total of three different composite compartments were manufactured for assembling the elevator. They are pcomp

81 000, pcomp 81 002 and pcomp 81 003 (figure 4). The lay-up of each panel compartment is listed in table 1. It is important to note that each panel consisted of two different materials: 5H carbon polyphenylene sulfide (PPS) fabric; and plain weave carbon PPS fabric. In table 1 the 5H carbon PPS fabric plies are indicated in bold, and the plain weave carbon PPS fabric plies are indicated using a normal font.

3.2. Steel ball impact experiment

The impact experiment was conducted on the elevator by using a steel ball. The diameter of the steel ball is 0.013. The weight of the steel ball is 8.40 g. The distance from the steel ball to the elevator surface was kept constant at 0.61 m for all of the impacts (figure 5). The impact energy of the steel sphere is 0.05 J. The only varying variable during this study was the locations of impact. There are three ribs in the elevator. Three impact locations were marked as red points on each rib (figure 6). A PAC Micro-30 sensor (manufactured by Mistras Group, Inc., Princeton Junction, NJ) was mounted to the spar of the elevator. Each impact location was repeatedly impacted 60 times. In total, 3600 impacts were conducted on ribs. AE signals were collected by the acquisition system during the experiment.

3.3. Acoustic emission setup

The AE acquisition system is manufactured by MISTRAS Group, Inc. (Princeton Junction, NJ). The amplitude threshold was set to 32 dB. The pre-trigger time was set to 256 μ s. This parameter ensures that the acquisition system does not miss the time attributed to signal initiation. The sampling rate was set to 2 MHz. The peak definition time (PDT), which refers to the time from threshold crossing to peak amplitude, was set to 200 μ s. This parameter is used to find the peak. It determines which peak can be applied to calculate rise time and amplitude when an AE signal has more than one peak. The hit definition time (HDT) was defined as 400 μ s in this study. This parameter determines when to stop the recording of a hit. The recording of the AE signal began when its voltage first exceeded the amplitude threshold. The recording is stopped at the moment that an amount of time equal to the HDT has passed without any threshold crossings. The HDT is typical twice the PDT. Lastly, the hit lockout time (HLT) was defined as 400 μ s. HLT is a timing parameter to assure that any threshold crossing that happens during this time will not be included in a hit waveform [39].

4. Model configuration selection

4.1. Configuration of random forest

A random forest classification model is utilized in this study for the purpose of source localization. This random forest contains 100 decision tree models. The Gini impurity of the node is used as the branching criterion. The input is a dataset of 3600 AE samples collected during the impact test. Each sample contains 15 features that are introduced in section 5.1. One hundred subsets are randomly selected from the original

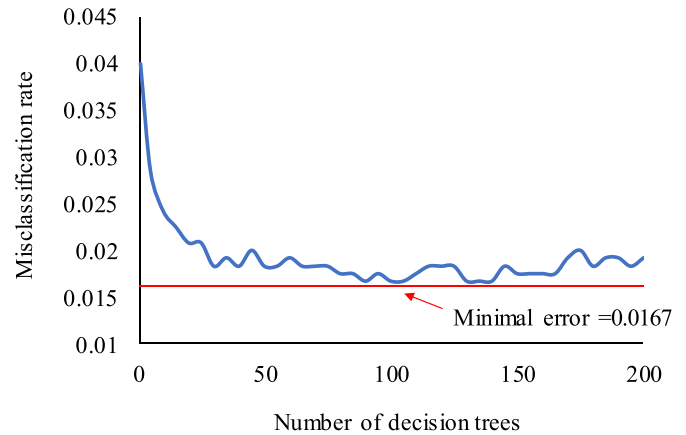


Figure 8. Misclassification rate.

input with replacements (bootstrapping) and then transferred to their own decision trees. The dimensions of each subset are the same as the original set, but only four features are randomly collected from 15 features. This is based on a criterion that the maximum number of features for a single decision tree can be the square root of the number of original features [40]. In the training process, the labels used by the decision trees are the zone number of the corresponding event. These decision trees work independently and give their own results. A final source localization result (zone number) is obtained by voting. Figure 7 shows the random forest algorithm applied in this paper.

The definition of the number of decision trees is shown in figure 8. Generally, with the increase of the base model, the random forest will converge to a lower generalization error [21]. When the number of base models exceeds a certain value, the error of the random forest basically converges. Continuing to increase the number of trees will not basically reduce the error. Moreover, it will slow the computing. Therefore, selecting an appropriate number is important. Trial-and-error testing was conducted by changing the number of decision trees from 1 to 200. The input data and corresponding labels are introduced in section 5.1. The misclassification rates versus the number of trees are plotted as the blue curve. It can be observed that the misclassification rates decreased from 0.04 to less than 0.02 when the tree number was increased to 30. The misclassification rates stay below 0.02 when continuing to increase the number to 200. The minimal rate is observed as 0.0167 when the number of trees is around 100–105, or 130–140. To optimize the random model for the fastest yet accurate purpose, the number of decision trees was defined as 100 in this study.

4.2. Topology of stacked autoencoder

The deep learning model utilized in this paper is an SAE composed of two autoencoders. Autoencoder 1 has a hidden layer with 100 neurons, and autoencoder 2 has a hidden layer with 50 neurons. Different from the random forest, the raw AE waveforms, and the corresponding Fast Fourier Transformation (FFT) magnitude of the impact events on the

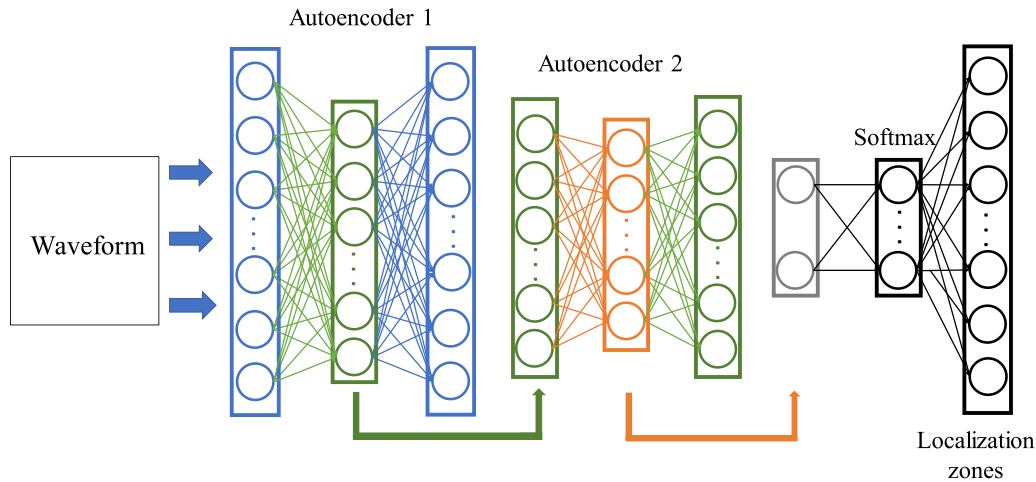


Figure 9. SAE impact source localization network.

Table 2. Descriptions of the input features for random forest.

| Features | Descriptions |
|----------------------------|--|
| Amplitude | The maximum amplitude at the peak |
| Count | The number of threshold crossings |
| Rise time | Time interval between first threshold crossing and maximum signal decibel |
| Duration | Time between the first and last signal crossing the threshold |
| Average frequency | Counts/duration |
| Root mean square (RMS) | The effective voltage with a characteristic time T_{RMS} for an average ranging between 10 and 1000 ms |
| Average signal level (ASL) | The effective voltage with a characteristic time T_{ASL} for an average ranging from 10 to 1000 ms |
| Energy | The measure of the electrical energy measured for an AE signal |
| Absolute energy | The absolute measure of the electrical energy measured for an AE signal |
| Peak frequency | Frequency of maximum signal contribution |
| Reverberation frequency | Frequency after the peak |
| Initial frequency | Frequency before the peak |
| Signal strength | A parameter to evaluate the AE source strength |
| Frequency centroid | A parameter to characterize the overall frequency content of an AE signal |
| Counts to peak (PCNTS) | The number of threshold crossings from the first threshold crossing to the peak |

elevator are used as the input data. During the training process, the labels of the inputs are the zone numbers that are attributed to each of the AE events. The source localization result is obtained as the output of this SAE neural network. The topology of the SAE used in this paper is shown in figure 9.

5. Results and discussion

5.1. Performance of random forest algorithm

5.1.1. Input preparation. The data is collected by the AE acquisition system during the steel ball impact test. Several features were extracted from the AE waveforms reconstructed by Whittaker–Shannon reconstruction. The features are used to reduce the amount of information carried by the signal to some specific values. They describe the characteristics of AE signals. In this study 15 main features were extracted from the original signals. The features and their descriptions are presented in table 2. Input samples which contain 3600 samples with all 15 features are adopted by the proposed random forest model.

5.1.2. Impact source localization. The previous investigation on which the current study is founded employed a BP ANN to localize the AE events of impacts. The elevator was divided into three zones based on the result of unsupervised pattern recognition [19]. The same division is utilized in this paper. The zonal divisions of the elevator are shown in figure 10. From left to right, the elevator is divided into zone 1, zone 2, and zone 3.

66.7% of the input samples were employed for training and the rest 33.3% were utilized for testing. The results are shown in figure 11(a) as a confusion matrix. As shown in the figure, 223 out of the 226 samples from zone 1 were correctly localized, while three samples were mistakenly assigned to zone 2. Out of the 474 samples of zone 2, 466 were correctly localized and eight were classified to zone 3 by error. Out of the 500 samples of zone 3, 491 were correctly localized, while nine samples were mistakenly classified to zone 2. According to figure 11(b), the localized recall of zones 1–3 are 98.7%, 98.3%, and 98.2%, respectively. The overall localized accuracy is 98.3%.

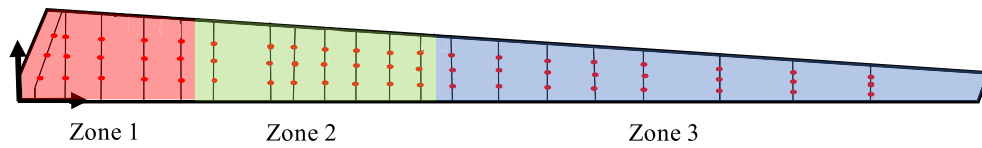


Figure 10. Zonal divisions.

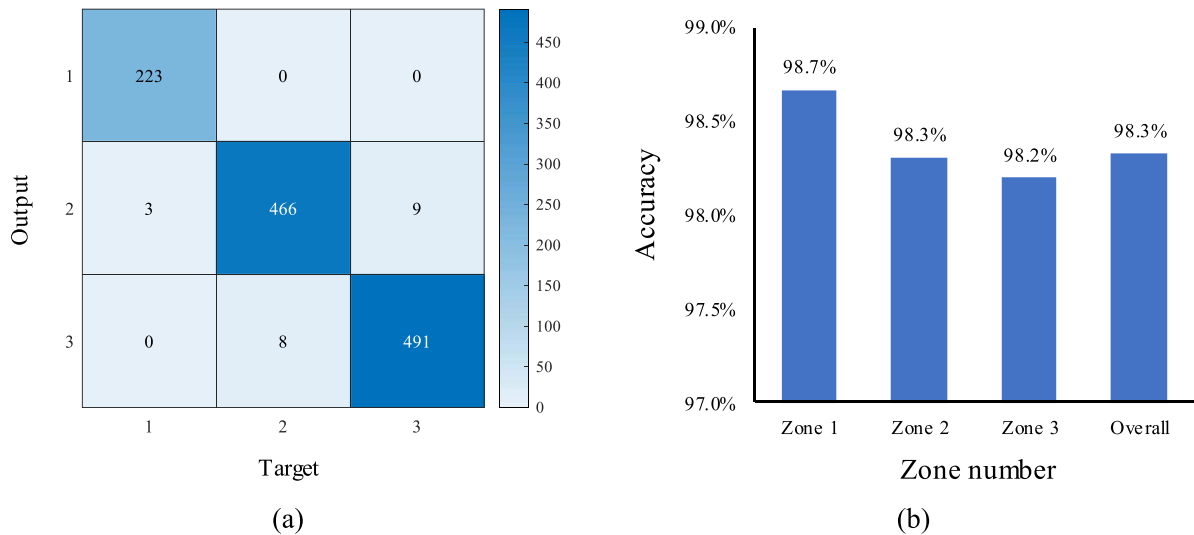


Figure 11. Performance of three zone classifications using random forest: (a) confusion matrix; and (b) accuracy of each zone.

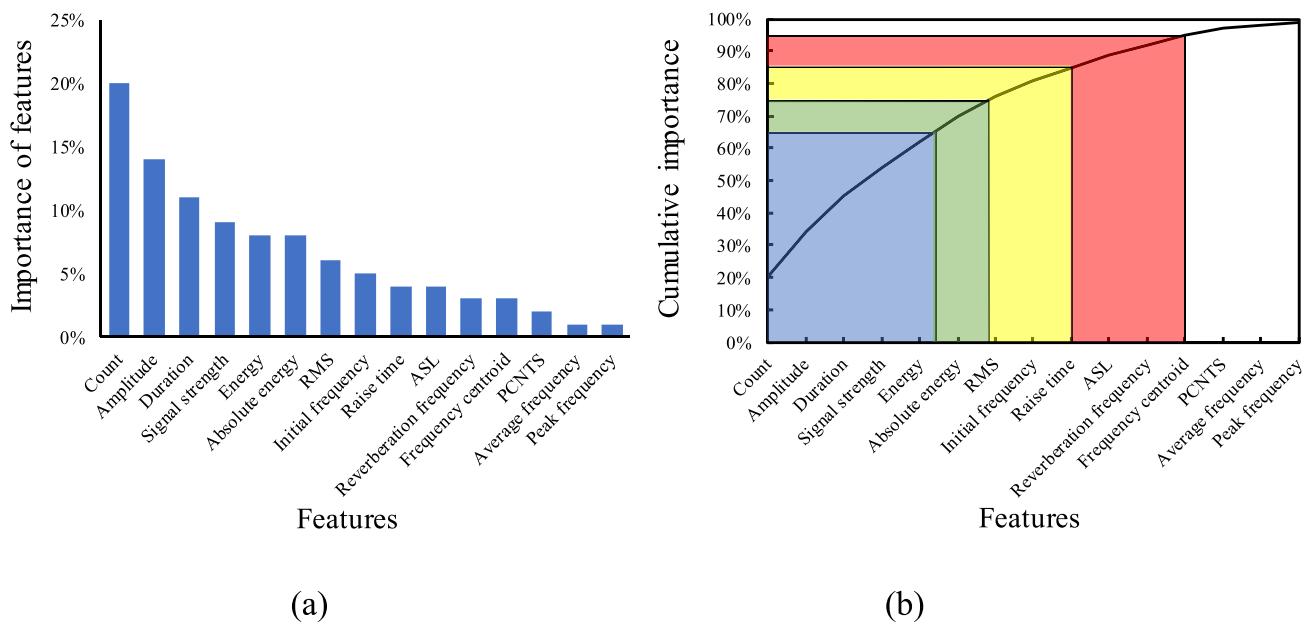


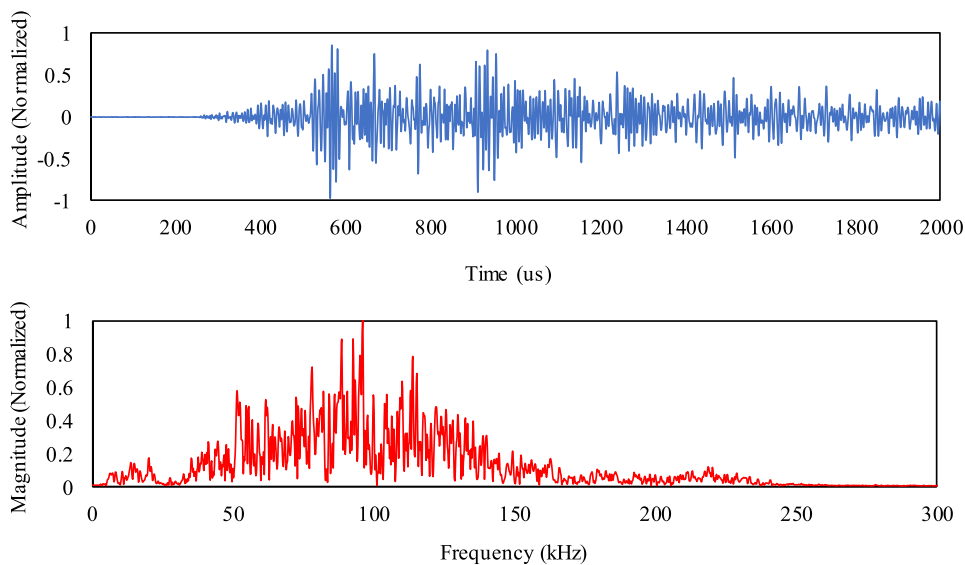
Figure 12. Importance of features: (a) ranking of importance, (b) cumulative importance.

5.1.3. Importance of features. According to the approach for calculating the importance of features, which was introduced in section 2.2, an important feature analysis was conducted. The importance of the 15 features is shown in figure 12(a) in descending order. The abscissa represents the different

feature names. The ordinate indicates the percentage of feature importance. According to figure 12(a), the ‘Count’, ‘Amplitude’, ‘Duration’ and ‘Signal strength’ of the features contribute to a large percentage of their overall importance, which have a significant impact on the localization results. Features

Table 3. Comparison of the performance by different input features.

| Input | Accuracy | Computing time (s) | Input storage (bytes) |
|------------------------------|----------|--------------------|-----------------------|
| Original input (15 features) | 98.3% | 0.3 | 299 008 |
| 95% importance (12 features) | 97.9% | 0.3 | 200 704 |
| 85% importance (9 features) | 97.8% | 0.2 | 163 840 |
| 75% importance (6 features) | 97.8% | 0.2 | 114 688 |
| 65% importance (5 features) | 97.8% | 0.2 | 90 112 |

**Figure 13.** Typical impact inputs: AE waveform (up), FFT magnitude (down).**Table 4.** Comparison of the accuracy by different input.

| Input | Accuracy |
|---------------|----------|
| AE waveform | 73.3% |
| FFT magnitude | 99.2% |

‘PCNTS’, ‘Average frequency’, and ‘Peak frequency’ have relatively low importance. Deleting them will not lead to a significant impact on the localization performance.

Based on the results above, a feature selection can be implemented for model optimization purposes. The computing time and the required input storage is reduced by reducing the input features, but accuracy is decreased. The optimum solution is to reduce as many features as possible while ensuring as high a performance as possible. The cumulative importance of features is plotted in figure 12(b). The blue region indicates that the first five features occupy 65% of the overall importance. The green region plus the blue region refer to the six features that have 75% of the importance. The yellow, green, and blue regions refer to the nine features with 85% of the importance. All of the 12 features within the regions with colors (blue, green, yellow, and red) occupy 95% of the overall importance. Four input samples with 12 features (Count, Amplitude, Duration, Signal strength, Energy, Absolute energy, RMS, Initial frequency, Raise time, ASL, Reverberation frequency, Frequency centroid), nine features

(Count, Amplitude, Duration, Signal strength, Energy, Absolute energy, RMS, Initial frequency, Raise time), six features (Count, Amplitude, Duration, Signal strength, Energy, Absolute energy) and five features (Count, Amplitude, Duration, Signal strength, Energy) were generated and assigned to the random forest model. Their localization accuracy, computing time, and required input storage were compared with the results of the original input. The detail is shown in table 3. It can be observed that reducing the input features from 15 to 5 slightly decreases the localization accuracy while the CPU computing time is reduced from 0.3 s to 0.2 s, and the required storage for input significantly decreases from 299 008 bytes to 90 112 bytes which means 69.9% of the input storage is saved.

5.2. Performance of the deep learning-based passive monitoring system

5.2.1. Input preparation.

In this study, the raw AE waveform and its FFT magnitude are considered as the input. A typical AE waveform of the impact event collected from the test and its FFT magnitude is presented in figure 13. In order to compare the performance of these two types of inputs, both the AE waveform and the FFT magnitude were assigned to the SAE network. The same labels for the three zones localization are utilized. 66.7% of the data was utilized as a training dataset, and the rest 33.3% was utilized as a testing dataset. The

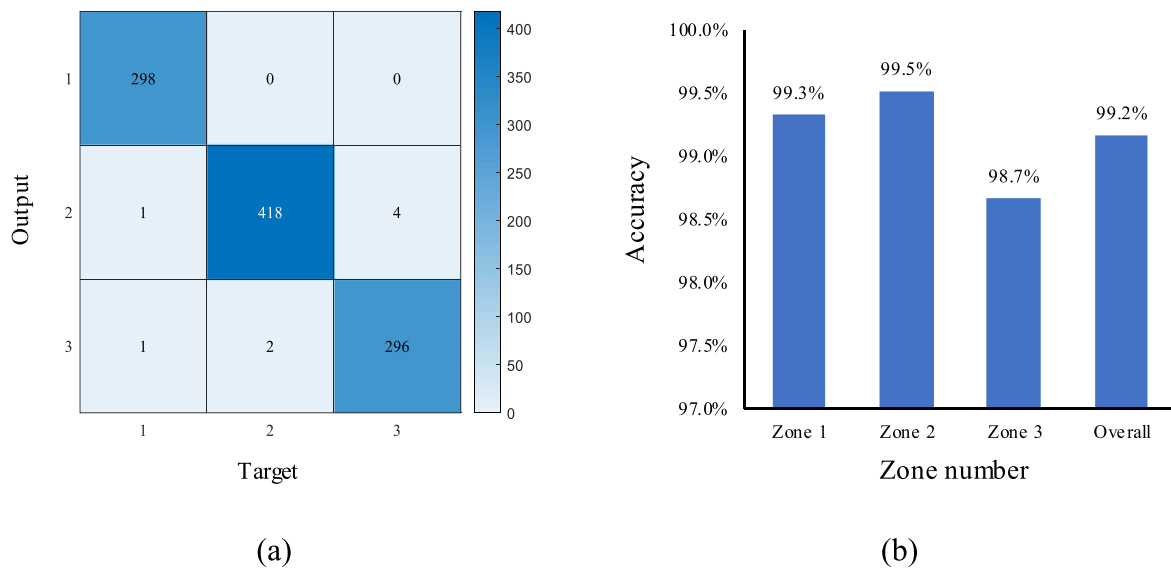


Figure 14. Performance of three zone classification using SAE. (a) Confusion matrix, (b) accuracy of each zone.

Table 5. Comparison of the performance.

| Input | Accuracy | Computing time (s) | Input storage (byte) |
|-----------------------------|----------|--------------------|----------------------|
| SAE | 99.2% | 681.4 | 458 219 520 |
| Random forest (15 features) | 98.3% | 0.3 | 299 008 |
| Random forest (12 features) | 97.9% | 0.3 | 200 704 |
| Random forest (9 features) | 97.8% | 0.2 | 163 840 |
| Random forest (6 features) | 97.8% | 0.2 | 114 688 |
| Random forest (5 features) | 97.8% | 0.2 | 90 112 |
| BP ANN | 96.0% | 0.8 | 299 008 |

optimum results of several times of training are as shown in table 4.

It can be observed that the SAE trained by the FFT magnitudes has an overall accuracy of 99.2%, while the SAE trained by the AE waveforms has an overall accuracy of 73.3%. The reason that the SAE trained by waveform obtained a low accuracy might be that the AE waveform contains low-frequency vibration information which is irrelevant to the AE source. When the feature extraction is carried out by autoencoders, the irrelevant vibration features will occupy some positions of the effective AE features, since the dimension of the feature set is a fixed number. This leads to the consequence that the final feature set for SoftMax classification contains unrelated vibration features which will greatly decrease the classification accuracy. The FFT magnitude contains multimodal and dispersive characteristics of AE and the low-frequency vibration is not significant in the FFT magnitude. The extracted feature set carried out by autoencoders has a high purity

of features that have a high correlation with the AE source, which leads to a better classification performance [41]. Therefore, the FFT magnitude was utilized as the input in this paper.

5.2.2. Impact source localization. The FFT magnitude dataset was adopted by the SAE network. The same divisions of zones as in the previous research [15] was applied. 66.7% of data were used as a training dataset, and 33.3% of data were used for testing dataset. The result is shown in figure 14 as a confusion matrix. According to figure 14(a), out of the 300 samples of zone 1, 298 were correctly localized, with one sample mistakenly assigned to zone 2, and another mistakenly assigned to zone 3. Out of the 450 samples of zone 2, 418 were correctly localized, and two samples were mistakenly assigned to zone 3. Out of the 300 samples of zone 3, 296 were correctly localized, while four samples were mistakenly assigned to zone 2. By considering figure 14(b), it can be observed that the localized recall of zones 1–3 is respectively 99.3%, 99.5%, and 98.7%. The overall localized accuracy is 99.2%.

5.3. Comparison of the proposed approaches

Based on the previous observation, the accuracy of the three-zone impact localization by using the BP ANN was 96.0% when the overlap of boundaries is not considered [20], which is lower than the accuracy of both random forest (98.3%) and SAE (99.2%). This indicates that SAE has the best source localization performance among the three methods. However, in practical applications, data storage size and computing time are very important evaluation factors. Because of the frequent impacts that may occur during flight, the acquisition system collects an enormous amount of AE signals. Minimizing the size of the data collected in the after-flight phase, computing time is important in optimizing the

passive monitoring system. The accuracy, computing time, and required input storage for the SAE, random forest, and BP ANN localization methods are shown in table 5. It can be observed that the SAE localization network in this study has the highest accuracy while it requires the longest computing time (681.4 s) and the largest input storage (458 219 520 bytes). The random forest significantly decreases the computing time (0.2 s) and input storage (90 112 bytes) when five features (Count, Amplitude, Signal strength, Energy) are utilized as an input, while it has an accuracy (97.8%) a little lower than SAE (99.2%) and higher than BP ANN (96.0%). Considering comprehensively, adopting a random forest localization method with optimized feature selection can greatly reduce the data storage size and obtain relatively high accuracy, which means the random forest-based passive monitoring system could be the optimum option for aircraft impact monitoring.

6. Conclusions

In this paper, AE-based passive monitoring approaches were investigated. An AE sensor was designed to be attached to an aircraft component during flight for the impact AE events collection in the in-flight phase. The collected data was utilized for impact source localization using random forest and SAE neural network in the simulated after-flight phase. An impact experiment was conducted on an aircraft elevator specimen to validate the efficiency of the proposed methods. A random forest model was built for AE feature selection and impact localization. An SAE neural network was also developed for source localization based on the raw AE signals. The main conclusions are as follows:

- (a) The SAE neural network has the best impact localization accuracy within SAE, BP ANN and random forest. However, it requires computing time and data storage much higher than both BP ANN, and random forest. Random forest with optimized feature selection led to an acceptable localization performance with minimal computing time and input data storage required.
- (b) More features lead to better localization performance of random forest, but the difference is not significant. Selecting the features with top importance led to an acceptable localization performance, meanwhile the computing time and required data storage were significantly reduced.
- (c) The FFT magnitudes are more appropriate to be employed as the input of the SAE neural network than AE waveforms because the AE waveform contains complex information which may reduce the localization performance.

The challenges in the application of the proposed methods in a real in-flight aircraft can be the interference of noise during flight. The difficulty of obtaining a sufficient number of labeled impact AE signals for training is also a challenge. Future research could focus on the study of the influence of noise and on generating training signals by numerical simulation. Another limitation of the proposed approaches in this

paper is that it only provides the large zonal source localization results. An investigation of the approach that can obtain the coordinates of the impact points are recommended as the subject of future research.

Data availability

The raw/processed data required to reproduce these findings can be made available upon request and with the written permission of the sponsor.

Acknowledgments


This paper was partially supported by GKN—Fokker Aerostructures BV, and the National Aeronautics and Space Administration (NASA) under the University Leadership Initiative program (Grant No. 80NSSC20M0165).

Conflicts of interest

The authors declare that there is no conflict of interest.

ORCID iDs

Li Ai  <https://orcid.org/0000-0003-2938-5533>

Mahmoud Bayat  <https://orcid.org/0000-0002-0990-7077>

References

- [1] Drury C G, Prabhu P and Gramopadhye A 1990 Task analysis of aircraft inspection activities: methods and findings *Proc. Human Factors Society Annual Meeting* vol 34 (Los Angeles, CA: SAGE Publications) pp 1181–5
- [2] Ivantsiv V, Spelt J K and Papini M 2009 Mass flow rate measurement in abrasive jets using acoustic emission *Meas. Sci. Technol.* **20** 095402
- [3] Austin R and Ziehl P 2011 High temperature health monitoring of organic matrix composites for aircraft engine applications *Proc. 8th Int. Workshop on Structural Health Monitoring*
- [4] Ono K 2011 Acoustic emission in materials research—a review *J. Acoust. Emission* **29** 284–308
- [5] Austin R, Forsyth D, Yu J, ElBatanouny M and Ziehl P 2014 Damage evaluation for high temperature CFRP components using acoustic emission monitoring *AIP Conf. Proc.* **1581** 501–5
- [6] Soltangharai V, Anay R, Hayes N W, Assi L, Le Pape Y, Ma Z J and Ziehl P 2018 Damage mechanism evaluation of large-scale concrete structures affected by alkali-silica reaction using acoustic emission *Appl. Sci.* **8** 2148
- [7] Ono K 2018 Review on structural health evaluation with acoustic emission *Appl. Sci.* **8** 958
- [8] Soltangharai V, Hill J W, Ai L, Anay R, Greer B, Bayat M and Ziehl P 2020 Acoustic emission technique to identify stress corrosion cracking damage *Struct. Eng. Mech.* **75** 723–36
- [9] Ai L, Greer B, Hill J, Soltangharai V, Anay R and Ziehl P 2019 Finite element modeling of acoustic emission in dry cask storage systems generated by cosine bell sources *AIP Conf. Proc.* **2102** 2

- [10] Ai L, Soltangharaei V, Anay R, Van Tooren M J and Ziehl P 2020 Data-driven source localization of impact on aircraft control surfaces *2020 IEEE Aerospace Conf. (7 March 2020)* (IEEE) pp 1–10
- [11] Soltangharaei V, Anay R, Ai L, Giannini E R, Zhu J and Ziehl P 2020 Temporal evaluation of ASR cracking in concrete specimens using acoustic emission *J. Mater. Civ. Eng.* **32** 04020285
- [12] Ono K, Mizutani Y and Takemoto M 2007 Analysis of acoustic emission from impact and fracture of CFRP laminates *J. Acoust. Emission* **25** 179–86
- [13] Han X, Zhao S, Cui X and Yan Y 2019 Localization of CO₂ gas leakages through acoustic emission multi-sensor fusion based on wavelet-RBFN modeling *Meas. Sci. Technol.* **30** 085007
- [14] Soltangharaei V, Ai L, Anay R, Bayat M and Ziehl P 2021 Implementation of information entropy, b-value, and regression analyses for temporal evaluation of acoustic emission data recorded during ASR cracking *Pract. Period. Struct. Des. Constr.* **26** 04020065
- [15] Santos-Leal E and Lopez R J 1995 Simultaneous measurement of acoustic emission and electrical resistance variation in stress-corrosion cracking *Meas. Sci. Technol.* **6** 188
- [16] Li D, Kuang K S and Koh C G 2017 Fatigue crack sizing in rail steel using crack closure-induced acoustic emission waves *Meas. Sci. Technol.* **28** 065601
- [17] Du S, Feng G, Li Z, Sarkodie-Gyan T, Wang J, Ma Z and Li W 2019 Measurement and prediction of granite damage evolution in deep mine seams using acoustic emission *Meas. Sci. Technol.* **30** 114002
- [18] Liu S, Li X, Li Z, Chen P, Yang X and Liu Y 2019 Energy distribution and fractal characterization of acoustic emission (AE) during coal deformation and fracturing *Measurement* **136** 122–31
- [19] Soltangharaei V, Anay R, Begrajka D, Bijman M, ElBatanouny M K, Ziehl P and Van Tooren M J 2019 A minimally invasive impact event detection system for aircraft movables *AIAA Scitech 2019 Forum* p 1268
- [20] Sexton R S, Dorsey R E and Johnson J D 1998 Toward global optimization of neural networks: a comparison of the genetic algorithm and backpropagation *Decis. Support Syst.* **22** 171–85
- [21] Bengio Y, Simard P and Frasconi P 1994 Learning long-term dependencies with gradient descent is difficult *IEEE Trans. Neural Netw.* **5** 157–66
- [22] Nourani V and Parhizkar M 2013 Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall–runoff modeling *J. Hydroinform.* **15** 829–48
- [23] Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
- [24] Shevchik S A, Saeidi F, Meylan B and Wasmer K 2016 Prediction of failure in lubricated surfaces using acoustic time–frequency features and random forest algorithm *IEEE Trans. Ind. Inform.* **13** 1541–53
- [25] Wang Z, Chegdani F, Yalamarti N, Takabi B, Tai B, El Mansori M and Bukkapatnam S 2020 Acoustic emission characterization of natural fiber reinforced plastic composite machining using a random forest machine learning model *J. Manuf. Sci. Eng.* **142** 031003
- [26] Iquebal A S, Pandagare S and Bukkapatnam S 2020 Learning acoustic emission signatures from a nanoindentation-based lithography process: towards rapid microstructure characterization *Tribol. Int.* **143** 106074
- [27] Bengio Y, Courville A C and Vincent P 2012 Unsupervised feature learning and deep learning: a review and new perspectives *CoRR* (arXiv:1206.5538) **1** 2012
- [28] Hinton G E and Salakhutdinov R R 2006 Reducing the dimensionality of data with neural networks *Science* **313** 504–7
- [29] Sadoughi M, Downey A, Bunge G, Ranawat A, Hu C and Laflamme S 2018 A deep learning-based approach for fault diagnosis of roller element bearings *Conf. Prognostics and Health Management Society (24 September 2018)* pp 1–7
- [30] Li C, Sanchez R V, Zurita G, Cerrada M, Cabrera D and Vásquez R E 2016 Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals *Mech. Syst. Signal Process.* **76** 283–93
- [31] He M and He D 2017 Deep learning based approach for bearing fault diagnosis *IEEE Trans. Ind. Appl.* **53** 3057–65
- [32] Shevchik S A, Kenel C, Leinenbach C and Wasmer K 2018 Acoustic emission for *in situ* quality monitoring in additive manufacturing using spectral convolutional neural networks *Addit. Manuf.* **21** 598–604
- [33] Ebrahimkhanlou A and Salamone S 2018 Single-sensor acoustic emission source localization in plate-like structures using deep learning *Aerospace* **5** 50
- [34] Raymer M G 1997 The Whittaker–Shannon sampling theorem for experimental reconstruction of free-space wave packets *J. Mod. Opt.* **44** 2565–74
- [35] Sandri M and Zuccolotto P 2006 Variable selection using random forests *Data Analysis, Classification and the Forward Search* (Berlin: Springer) pp 263–70
- [36] Ng A 2011 Sparse autoencoder *CS294A Lect. Notes* **72** 1–9
- [37] Bengio Y, Lamblin P, Popovici D and Larochelle H 2007 Greedy layer-wise training of deep networks *Advances in Neural Information Processing Systems* (Cambridge, MA: MIT Press) pp 153–60
- [38] Sun R, Chen Y, Dubey A and Pugliese P 2020 Hybrid electric buses fuel consumption prediction based on real-world driving data *Transp. Res. D* **91** 102637
- [39] Laksimi A, Benmedakhene S and Bounouas L 1999 Monitoring acoustic emission during tensile loading of thermoplastic composites materials *Proc. ICCM (5 July)* vol 12
- [40] Genuer R, Poggi J M and Tuleau-Malot C 2010 Variable selection using random forests *Pattern Recognit. Lett.* **31** 2225–36
- [41] Ebrahimkhanlou A, Dubuc B and Salamone S 2019 A generalizable deep learning framework for localizing and characterizing acoustic emission sources in riveted metallic panels *Mech. Syst. Signal Process.* **130** 248–72