Special Issue Article

Structural Health Monitoring

Structural Health Monitoring
2021, Vol. 20(4) 1597–1608

DOI: 10.1177/1475921720923147
journals.sagepub.com/home/shm
SAGE

If structure can exclaim: a novel robotic-assisted percussion method for spatial bolt-ball joint looseness detection

Furui Wang[1], Aryan Mobiny[2], Hien Van Nguyen[2] and Gangbing Song[1]

Abstract

In proportion to the immense construction of spatial structures is the emergence of catastrophes related to structural damages (e.g. loose connections), thus rendering personal injury and property loss. It is therefore essential to detect spatial bolt looseness. Current methods for detecting spatial bolt looseness mostly focus on contact-type measurement, which may not be practical in some cases. Thus, inspired by the sound-based human diagnostic approach, we develop a novel percussion method using the Mel-frequency cepstral coefficient and the memory-augmented neural network in this article. In comparison with current investigations, the main contribution of this article is the detection of multi-bolt looseness for the first time with higher accuracy than prior methods. In particular, in terms of new data obtained via similar joints, the memory-augmented neural network can help avoid inefficient relearn and assimilate new data to provide accurate prediction with only a few data samples, which effectively improves the robustness of detection. Furthermore, percussion was implemented with a robotic arm instead of manual operation, which preliminarily explores the potential of implementing automation applications in real industries. Finally, experimental results demonstrate the effectiveness of the proposed method, which can guide future development of cyber-physics systems for structural health detection.

Keywords

Bolt looseness detection, percussion method, structural health monitoring, Mel-frequency cepstral coefficient, memory-augmented neural network

Introduction

Due to its advantages including beauteous appearance, lightweight, and a sense of transparency, the spatial structure has been widely used across different infrastructures such as airports and sports stadiums. Generally, spatial structures are composed of many members and joints, which can hold members together and support loading. Compared to other joint systems (e.g. the welded hollow spherical joints[1] and the space-truss connectors),[2] the bolt-ball joints[3] are preferred due to merits of being easy-to-implement and low material consumption. For instance, the welded connections are welded on-site, which requires extra labor and capital. However, prior investigation[4] has proven that bolted joint stiffness is an essential factor in ensuring structural integrity since it affects the stability behavior and load-bearing capacity of the whole spatial structure significantly. Thus, some numerical[5,6] and experimental[7,8] studies have been developed to research the mechanical behavior of bolt-ball joints. Nevertheless, owing to large body mass and long service life, we may encounter accidents of spatial structures, caused by different natural disasters such as hurricanes and earthquakes. Moreover, since the bending and shear forces are transferred between each member and ball node through only one bolt, the bolt looseness can lead to stiffness deterioration and stress redistribution that collapse the whole structure. Even though some new joints[9,10] have been designed and fabricated to alleviate this issue, we still need some effective methods to detect the degradation and failures of bolt-ball joints, particularly bolt looseness.

[1]Department of Mechanical Engineering, University of Houston, Houston, TX, USA
[2]Department of Electrical & Computer Engineering, University of Houston, Houston, TX, USA

Corresponding author:
Gangbing Song, Department of Mechanical Engineering, University of Houston, Houston, TX 77204, USA.
Email: gsong@uh.edu

In the past decades, different structural health monitoring (SHM) methods have been employed to detect damages of bolt-ball joints, including the vibration-based method combined with numerous signal processing strategies such as empirical mode decomposition (EMD),[11] rank-revealing QR (RRQR) decomposition,[12] fuzzy c-means (FCM) clustering algorithm,[13] and continuous wavelet transform (CWT) method.[14] Subsequently, with the rapid development of smart materials and structures (SMS), the SMS-enabled SHM methods have attracted more attention in detecting looseness of various bolts, for example, the vibration-based method,[15,16] the active sensing method,[17–19] the electro-mechanical impedance (EMI) method,[20,21] the nonlinear ultrasound method,[22–27] and the eddy current method.[28] Particularly, Xu et al.[29] employed active sensing and EMI methods[30] to detect the looseness of spatial bolt-ball joints, which demonstrated excellent performance. However, instead of multiple loosening, Xu et al.[30] focus on detecting single-bar looseness of the spatial bolt-ball joint, which dramatically reduces the practical value-in-use. Moreover, all the SHM methods mentioned above depend on sensor deployment; in other words, constant contact between sensors and structures is required. This deployment can incur high costs and is impractical in some complex cases. To address this issue, some machine-vision-based methods[31] are proposed recently to avoid contact detection of bolt looseness. However, these methods have some intrinsic drawbacks; for instance, the inception of looseness and some exceptional cases (e.g. the loose bolt that has one circle rotation exactly) cannot be identified effectively. Therefore, we still need a more practical method to detect looseness in spatial bolt-ball joints.

Since ancient times, the sound has been a diagnostic indicator for human health; for instance, infant cries can express hunger, pain, drowsiness, and asphyxia. It is therefore natural that we can expect the potential "exclaim" of structures (a.k.a. percussion-induced sound signals) to characterize the structural performance. The percussion-based method has been brought into focus in detecting the health status of multiple structures[32–34] recently since radiation acoustic signals can correspond to structural damages that induce changes in mechanical properties. Particularly, several investigations[35–38] have been conducted to detect bolt looseness via the percussion-based methods. However, the above methods depend on manual tapping, which might be impractical for spatial bolt-ball joints. Furthermore, they all only consider single bolt looseness, whose real applications are rare in industries. On the contrary, the robotic-assisted percussion method is expected to be highly potent in detecting looseness of spatial bolt-ball joints, since several robotic-assisted SHM methods[39,40] have proven their superiority by virtue of being portable and low-cost. As illustrated in Figure 1, a climbing robot, which is equipped with a hammer and a microphone, can crawl among bolt-ball joints to detect their looseness by tapping and processing the percussion-induced sound signals.

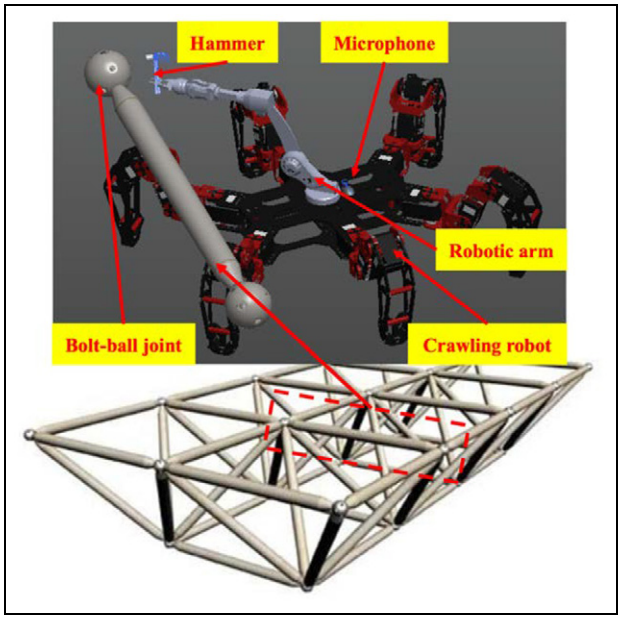

Figure 1. Schematic of robotic-assisted detection of spatial bolt-ball joint looseness.

The main contribution of this article is the detection of multiple looseness of the spatial bolt-ball joint for the first time by developing a new robotic-assisted percussion method, which requires no contact sensor deployment. In addition, we achieve higher identification accuracy than prior percussion-based methods that focus on single bolt looseness. The proposed method employs the Mel-frequency cepstral coefficient (MFCC) algorithm to extract features from the percussion-induced sound signals or "exclaim" of structures. The other main contribution is the application of memory-augmented neural network (MANN) to classify different MFCC feature sets. In comparison with the current convolutional neural networks (CNNs), the MANN can achieve better performance, particularly when it encounters new data. In other words, the proposed method can construct classification for similar applications with minimal prior knowledge and can, therefore, help relieve the pressure from feature extractors. Another advantage of MANN over CNN is better adaptability, guaranteed through the characteristics of the MANN. Finally, multiple proof-of-concept experiments were conducted to verify the effectiveness of the proposed method. The rest of this article is organized as follows. An overview of related literature is presented in section "An overview of related literature."

The theoretical background of the MFCC and the MANN is introduced in section "Methods." Section "Experimental setup" describes the experimental apparatus in detail. The results are discussed in section "Results and discussion," and section "Conclusion" concludes this article. Overall, the proposed method opens up a new frontier for further investigations on the detection of spatial bolt-ball joint loosening.

An overview of related literature

In this section, we conduct a detailed literature review to better illustrate the research status of current percussion-based methods for bolt looseness detection and CNN-based MFCC feature classification. Through this review, we can fully confirm the main contribution of the proposed method, which can be found in the following subsection in detail.

Percussion-based methods for bolt looseness detection

The principle of the percussion method has been widely employed in our daily life to diagnose structural integrity. For example, we tap melons to determine their ripeness by listening to the sound, and the integrity of vitreous structures can be judged according to the tone (namely, the bright or dull sound) when we knock on them. Moreover, the pleural effusion associated with an unhealthy lung can be diagnosed by tapping the patient's chest and listening sound. Overall, the principle of the percussion-based method is that the changes in mechanical properties, which are caused by structural damages, correspond to generated acoustic signals when tapping on structures.[41]

Recently, the superiority of the percussion method for bolt looseness detection has been noticed. Wang et al.[35] developed analytical modeling and numerical simulation of radiation acoustic signals caused by percussion on a bolted connection, which laid a solid foundation for future investigations. However, it is well known that analytical models, which are time-consuming, are not suitable for a quick inspection. Subsequently, some more practical methods[36–38] were developed to achieve bolt looseness detection more efficiently, while they all have some demerits. For instance, the feature extraction (power spectral density, PSD) of sound signals in Kong's method[36] was conducted manually, thus incurring inaccuracy due to subjectivity. Meanwhile, Zhang's method[37] employed the MFCC algorithm to extract features from sound signals, which requires no manual selection. However, only the first order of the MFCC matrix is fed into a support vector machine (SVM) to achieve the classification. This implementation may be questionable since other orders of the MFCC matrix also contribute to the features of percussion-induced sound signals. Yuan et al.[38] extracted the intrinsic multiscale entropy (IME) of sound signals as features and trained a backpropagation neural network (BPNN) model to achieve classification of various bolt preloads. However, all these three methods focus on single-bolt connections and may be incapable of handling multi-bolt connections; thus, we still need a new method to detect multi-bolt looseness. In addition, current percussion methods are all implemented manually, which lack the potential of automatic applications.

MFCC feature classification using CNN-based methods

It is worth noting that new opportunities, which can break through the current bottleneck of one research filed, are always generated via migrating approaches from another research field. That is to say, we can expect the potential of automatic speech recognition (ASR) techniques to process percussion-induced sound signals for bolt looseness identification since these signals are similar to human speech, which has rich characters to be extracted. Thus, Zhang's method[37] provides good inspiration for our research; however, a new classification method is required to process the MFCC feature sets to detect bolt looseness. Recently, the deep learning technology has attracted a lot of attention across multiple applications. Particularly, the CNNs[42,43] have proven their better performance over current ASR methods, including the hidden Markov models (HMMs), the Gaussian mixture models (GMMs), and their combination (i.e. the GMM-HMMs). This improvement can be partially attributed to the unique capacity of the CNNs in characterizing complex functions[44] since the CNNs can automatically integrate different levels (from low to high) of features in an end-to-end classification architecture. Moreover, the combination of CNN and MFCC has demonstrated its superiority in speech recognition and other applications of pattern recognition. For instance, Liu et al.[45] proposed a CNN-MFCC hybrid method for short utterance speaker recognition, and Phan et al.[46] developed a Label-Tree Embeddings (LTE) algorithm, which employs the CNNs to classify different labeled low-level features including the MFCC, to achieve audio scene classification. Similar investigations have also been reported in previous literature.[47–52] Even though these studies[53–60] demonstrate that CNNs have significant capacity to process the MFCC and capture features for classification, some demerits limit the performance. Notably, current CNN-based methods

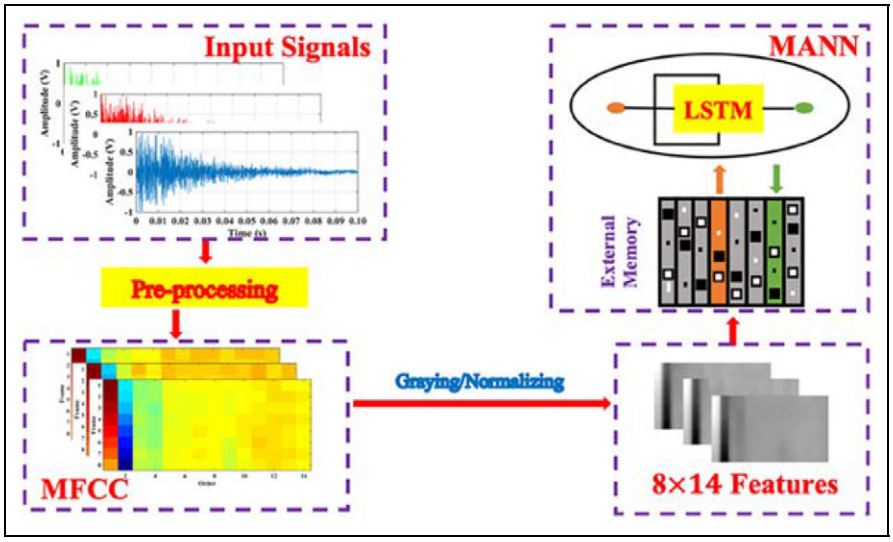

Figure 2. Flowchart of the proposed method.

require large amounts of data to train the models extensively, and the models need to relearn their inherent parameters to incorporate classifications when encountering new input. On the other hand, we always expect to achieve rapid inference with smaller quantities of data for most issues of interest. Therefore, we apply the MANN to process MFCC features in this article to obviate the downsides of current methods. More details of MANN are presented in section "MANN."

Methods

The flowchart of the proposed method in this article is illustrated in Figure 2. First, the percussion-induced sound signals are pre-processed and then converted to a representation in the time-frequency domain through the MFCC. After graying and normalizing the MFCC features, we fed them into the MANN, which consists of a long short-term memory (LSTM) controller and an external memory. More details of the MFCC and the MANN will be introduced in the following subsection.

MFCC

To reduce dimensionality and quantify the signal probability for speech recognition, we can project the original acoustic signal to a feature space through many feature extraction methods. One of the most popular approaches is the MFCC,[53] which represents the nonlinear cepstrum of signals, and its computation process is depicted in Figure 3. After splitting and windowing the initial input signal in the time domain to prevent frequency distortion, we apply the discrete Fourier transform (DFT) to obtain the power spectrum in the frequency domain. Subsequently, we convert the frequency (Hertz scale) to the Mel scale through a filter bank, which can also achieve frequency filtering. Finally, by utilizing the discrete cosine transform (DCT) on the Mel power spectrum (in logarithm), we extract the MFCC features. The detailed process is given as follows:

(a) After applying a Hamming window to segment the initial signal $x(n)$, we can realize the DFT through

$$x(k) = \sum_{n=0}^{M_s-1} x(n) e^{-j2\pi nm/M_s} \tag{1}$$

where $0 \leqslant k \leqslant M_s$, and M_s is the total points for the DFT; k is the parameter for M_s.

(b) Typically, the Mel-scale filter bank consists of several triangular filters, whose distribution is based on the Mel-scaling estimation, and the corresponding response $H_i(k)$ of ith filter takes the form of

$$H_i(k) = \begin{cases} 0 & k < k_{b_i-1} \\ \dfrac{k - k_{b_i-1}}{k_{b_i} - k_{b_i-1}} & k_{b_i-1} \leqslant k \leqslant k_{b_i} \\ \dfrac{k_{b_i+1} - k}{k_{b_i+1} - k_{b_i}} & k_{b_i} \leqslant k \leqslant k_{b_i+1} \\ 0 & k > k_{b_i+1} \end{cases}$$

$$k_{b_i} = (M_s/F_s) f_{mel}^{-1}[f_{mel}(f_{\min}) + i\{f_{mel}(f_{\max}) - f_{mel}(f_{\min})\}/(Q+1)] \tag{2}$$

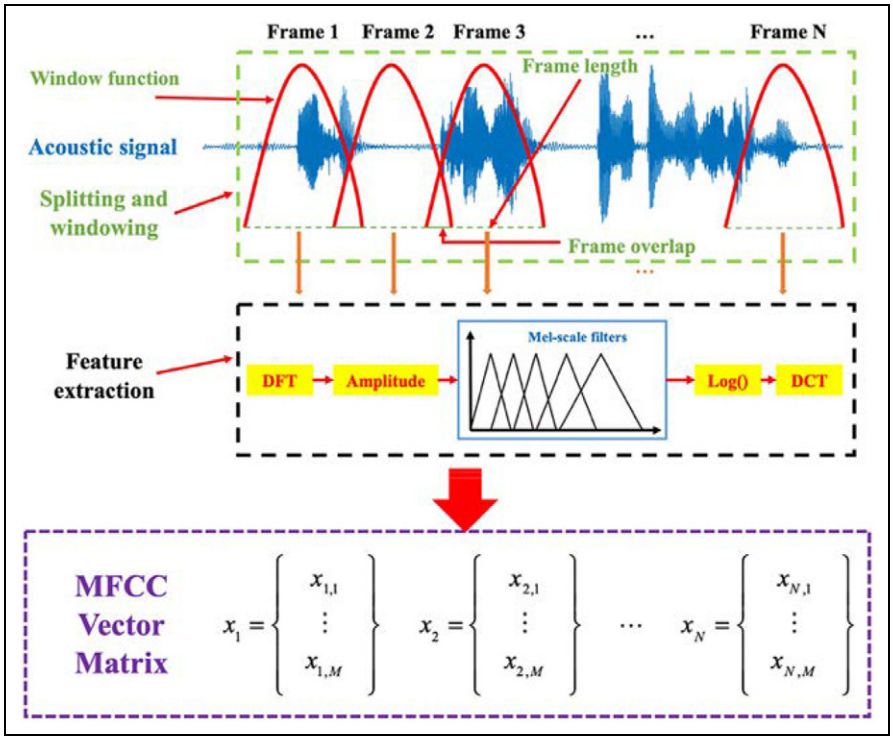

Figure 3. Flowchart of the MFCC feature extraction and computation process.

where k_{b_i} is the boundary points of the filter,[53] Q is the total of filters, and $f_{\max}$ and $f_{\min}$ are the maximum and minimum of the frequency range. Then, based on O'Shaughnessy's theory,[55] the function f_{mel} and its inverse f_{mel}^{-1} can be obtained as follows

$$\begin{cases} f_{mel} = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right) \\ f_{mel}^{-1} = 700\left(10^{f_{mel}/2595} - 1\right) \end{cases} \tag{3}$$

(c) Finally, we compute the MFCC coefficients $c(n)$ by using the DCT to process the energy spectrum $s(i)$ of filter bank

$$\begin{cases} c(n) = \sum_{i=0}^{Q-1} s(i) \cos\left(\frac{n(i-0.5)}{Q}\pi\right) \\ s(i) = \ln\left[\sum_{k=0}^{Q-1} |x(k)|^2 H_i(k)\right] \end{cases} \tag{4}$$

MANN

As we have discussed earlier, an especially daunting challenge we face with deep learning is the generation of new behavior (e.g. classification) according to inference from a few scraps of information.[61] This requirement has gone beyond the capacity of current machine intelligence to some extent. Contemporary gradient-based solutions must re-learn parameters completely from the new data, and when there are little data, these strategies are prone to poor learning, incurring catastrophic interference. Several attempts have been made at alleviating this obstacle, and some solutions[62] have proven their potential with rapid learning based on sparse data with the notion of meta-learning. Moreover, Neural Turing Machines (NTMs)[63] and memory networks[64] were proposed to meet the requirement better. Particularly, Santoro et al.[65] developed the MANN, which has a better capacity to accomplish meta-learning.[66] Subsequently, several investigations[67–70] have proven the advantages of MANN.

The architecture of MANN is illustrated in Figure 4. MANN consists of several components: a controller (generally, we employ the LSTM), an output distribution, an external memory, and read/write heads. After receiving the input data (x_t, y_{t-1}), the controller can update the state as

$$\hat{g}^f, \hat{g}^i, \hat{g}^o, \hat{\mathbf{u}} = \mathbf{W}^{xh}(x_t, y_{t-1}) + \mathbf{W}^{hh}\mathbf{h}_{t-1} + \mathbf{b}^h \tag{5}$$

where $\hat{g}^f, \hat{g}^i, \hat{g}^o$ are the gates of forget, input, and output, respectively; $\mathbf{W}^{xh}(x_t, y_{t-1})$ is the weight from the input (x_t, y_{t-1}) to the hidden state; $\mathbf{W}^{hh}$ is the weight between two different hidden states; $\mathbf{h}_{t-1}$ is the hidden state under data label y_{t-1}; and $\mathbf{b}^h$ is the bias of the hidden state.

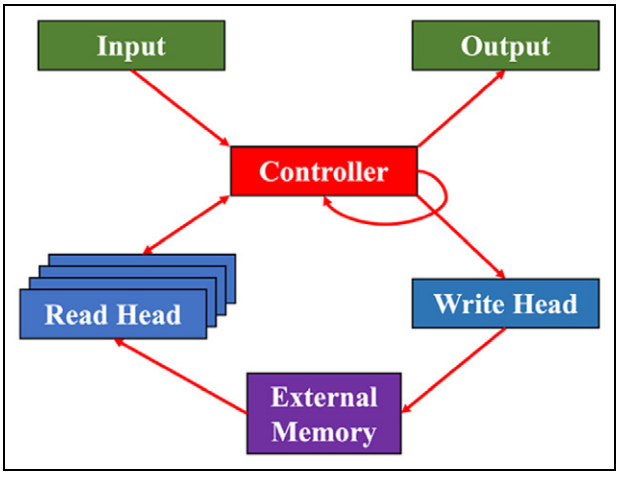

Figure 4. The architecture of the MANN.[65]

Subsequently, the concatenated output of the controller can be expressed as

$$\mathbf{o}_t = (\mathbf{h}_t, \mathbf{r}_t) \tag{6}$$

where $\mathbf{h}_t = \sigma(\hat{g}^o) \odot \tanh(\mathbf{c}_t)$; $\mathbf{c}_t = \sigma(\hat{g}^f) \odot \mathbf{c}_{t-1} + \sigma(\hat{g}^i) \odot \tanh(\hat{\mathbf{u}})$ is the cell state; $\mathbf{r}_t$ is the read vector according to the external memory $\mathbf{M}_t$; and $\sigma()$ is the sigmoid function. It is worth noting that the read vector $\mathbf{r}_t$ is retrieved via read-weight vector w_t^r as

$$\begin{aligned} \mathbf{r}_t &\leftarrow \sum_i w_t^r(i)\mathbf{M}_t(i) \\ w_t^r(i) &\leftarrow \frac{\exp(K(k_t, \mathbf{M}_t(i)))}{\sum_j \exp(K(k_t, \mathbf{M}_t(j)))} \end{aligned} \tag{7}$$

where $K(k_t, \mathbf{M}_t(i)) = k_t \cdot \mathbf{M}_t(i) / ||k_t|| ||\mathbf{M}_t(i)||$ is the cosine distance between the query key vector and each row of $\mathbf{M}_t$.

Similarly, Least Recently Used Access (LRUA) is employed to implement memory write function, and more details can be found in prior investigation.[65]

Finally, the output distribution can be calculated via $\mathbf{o}_t$, producing the classification probability $\mathbf{p}_t$ as

$$\mathbf{p}_t(i) \leftarrow \frac{\exp(\mathbf{W}^{op}(i)\mathbf{o}_t)}{\sum_j \exp(\mathbf{W}^{op}(i)\mathbf{o}_t)} \tag{8}$$

where $\mathbf{W}^{op}$ is weight of the output, and the episode loss $\mathcal{L}(\theta)$ can be obtained through

$$\mathcal{L}(\theta) = -\sum_t y_t^T \log \mathbf{p}_t \tag{9}$$

Experimental setup

To prove the capacity and verify the effectiveness of the proposed method, we conducted multiple proof-of-concept experiments on a 6-bay, 83-member spatial bolt-ball joint structure (size: 0.35 m length, 0.35 m width, and 2.1 m height), which is made of aluminum. As depicted in Figure 5, the experimental apparatus consists of a computer, an acoustic signal acquisition interface (Scarlett 18i8; Focusrite), a digital torque wrench, a robotic arm with control handle (OWI-535; OWI Robotic), a microphone (Ambeo VR; Sennheiser) held through a tripod, and a designed hammer. It is worth noting that a copper ball and a steel spring constitute the hammer since this design can better simulate the action of "tapping," rather than the "peening" when we apply the robotic arm that has no spring-back (see Supplementary Material). The tapping point is on the ball joint, while no accurate position can be guaranteed due to the spring. In other words, the copper ball does not need to hit the same spot in the repeated tapping. Generally, this scenario is preferred since it is more conducive to industrial applications. Moreover, the classification performance, which will be discussed in the next section, can demonstrate that this random tapping position has no significant influence on the damage detection results. The microphone has a pre-selected distance from the impact point (about 0.2 m) to reduce the effect of airflow, and its sampling time and sampling rate are set to 0.1 s and 48 kHz, respectively.

In this article, we considered multiple scenarios, including single-bar looseness and multi-bar looseness of the spatial bolt-ball joint. Furthermore, to demonstrate the superiority of the MANN, we tapped another bolt-ball joint via the robotic arm to conduct a new training and testing dataset, that is to say, the training dataset and the testing dataset for Case 7 are from different joints. More details of experimental arrangements are summarized in Table 1. For each case (from Case 1 to Case 6), we repeated the percussion tests 100 times to construct the dataset, in which the training set and testing set include 80% and 20% of the dataset, respectively. On the contrary, in terms of Case 7, the testing set has 20 samples, and the number of training sets is set to 2, 4, 6, 8, and 10, respectively.

Results and discussion

After down-mixing the percussion-induced sound signals to the mono channel (taking mean values from the left and right channels), we depicted samples under Cases 1–6 in Figure 6. Then, we employed the MFCC to process the pretreated signals to obtain feature sets, whose sample images under each case are also given in Figure 6. Typically, the size of samples affects the classification performance significantly: oversized samples increase the complexity and lead to redundant computation, while sufficient information cannot be presented

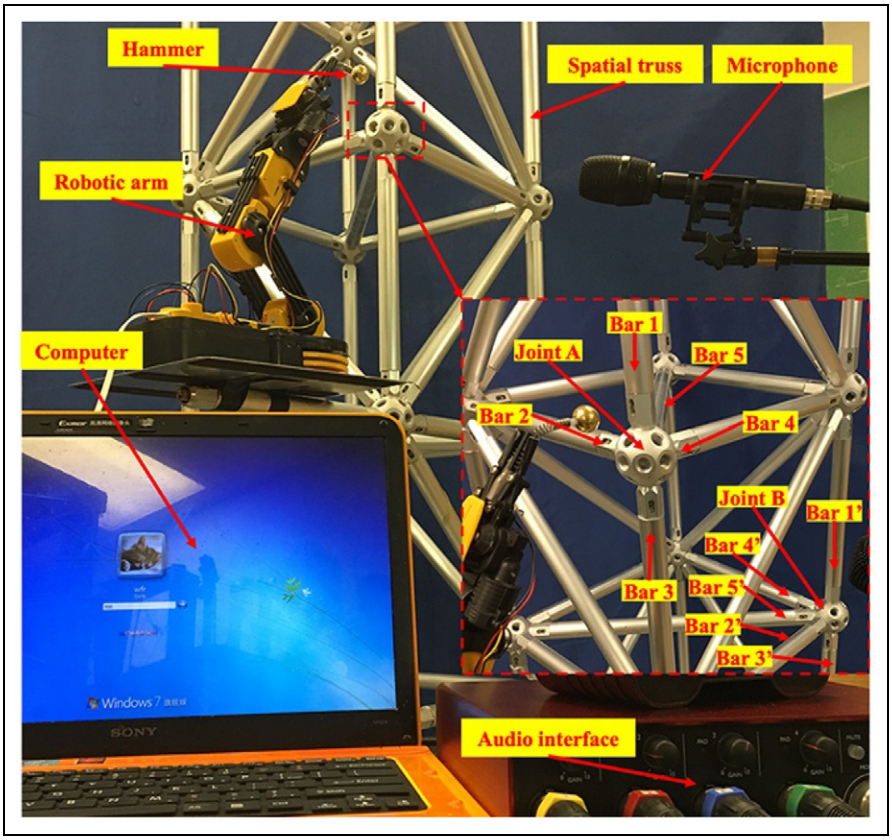

Figure 5. Experimental setup.

Table 1. Details of different experimental scenarios.

Case	Joint	Looseness	Tightened (20 N m)
1	A	N/A	Bars 1, 2, 3, 4, and 5
2		Bar 1 (10 N m)	Bars 2, 3, 4, and 5
3		Bar 1 (0 N m)	Bars 2, 3, 4, and 5
4		Bars 1 and 2 (0 N m)	Bars 3, 4, and 5
5		Bars 1, 2, and 3 (0 N m)	Bars 4 and 5
6		Bars 1, 2, 3, and 4 (0 N m)	Bar 5
7	B	Corresponding to Cases 1–6	

through too small samples adequately. Therefore, we conducted the trial and error, and the frame length was set as 10 ms with an overlap of 5 ms, thus rendering the size of the MFCC (8×14).

After converting the MFCC matrixes to the grayscale maps, we normalized the features under each case by subtracting mean and dividing them with standard deviation, which was obtained over the whole dataset. Then, the normalized MFCC features under each case were fed into the MANN model, which was conducted and achieved on the TensorFlow framework. Meanwhile, the data augmentation was implemented through random translation and rotation of inputs. To train the MANN, we first sorted the MFCC images with corresponding five-step labels, and each sequence was regarded as an episode. The ADAM (adaptive moment estimation) optimizer with a minibatch size of 32, whose configuration is the same as the CNN, was employed to construct the MANN, and a grid search was implemented to figure out the best values of parameters. In this article, 128 memory slots with a size of 40 were found to be the optimal solution, and the size of the LSTM controller was 200. In addition, the learning rate was set to $1e^{-4}$, and the number of reads from memory is 4 (with write decay of 0.99). Moreover, to enhance the computation efficiency, we employed a

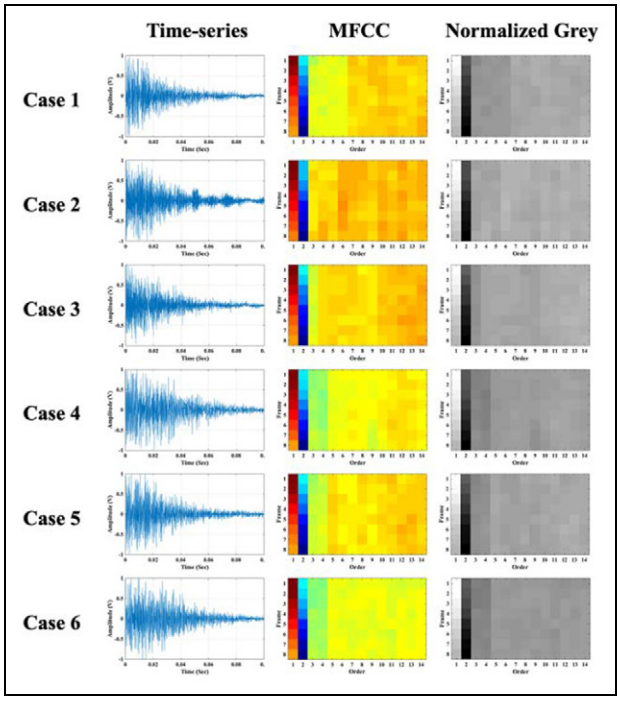

Figure 6. Samples of sound signal and MFCC features under different cases.

GPU (graphics processing unit, Nvidia GTX 960) without the requirement of GPU programming since the TensorFlow framework provides the support. Subsequently, to verify the effectiveness of the MANN, we computed its classification accuracies with maximum episodes of 100,000, as depicted in Figure 7. It is worth noting that these accuracies are computed for up to 10 feedbacks. That is to say, the instance accuracy in Figure 7 means the classification accuracy of the corresponding observation order of the whole cases. According to Figure 7, we find that the MANN achieves the highest testing accuracy after 25,000 episodes with 100% through the 10th feedback. Moreover, to further confirm the superiority of the MANN, we also implemented current percussion-based bolt looseness inspection methods[36–38] and the CNN-based MFCC classification, and the comparison results are presented in Table 2. It can be found that the CNN-based method outperforms the SVM, BPNN, and the decision tree in classifying the audio signals, while the proposed MANN has the best performance, which indicates that the proposed method possesses a stronger capacity for multi-bolt looseness detection.

Subsequently, we test the anti-noising capacity of the proposed method. Since the proposed method is still in the initial exploration stage (i.e. it is temporarily difficult to implement on-site experiment to realize the actual environmental noise), we add different white Gaussian noises into the testing data to obtain

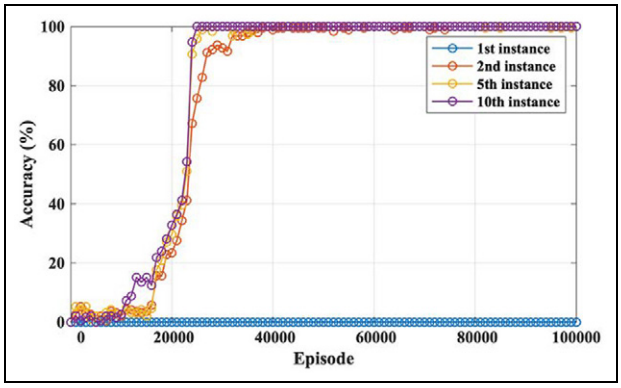

Figure 7. Testing accuracies under different cases (1–6) using the MANN.

Table 2. Comparison of testing classification accuracies between the proposed method and current methods.

Model	Instance (% accuracy)
PSD + DT[36]	66.67
MFCC + SVM[37]	92.17
IME + BPNN[38]	81.33
MFCC + CNN	99.17
MFCC + MANN	**100.00**

PSD: power spectral density; DT: decision tree; MFCC: Mel-frequency cepstral coefficient; SVM: support vector machine; IME: intrinsic multiscale entropy; BPNN: backpropagation neural network; CNN: convolutional neural network; MANN: memory-augmented neural network. Bold values are used to highlight the classification accuracy of the proposed method.

corresponding signal-to-noise ratio (SNR), which can be defined as

$$\mathrm{SNR}_{dB} = 10\log_{10}\left(\frac{A_{signal}^2}{A_{noise}^2}\right) \tag{10}$$

where A_{signal} and A_{noise} are the amplitude of the signal and noise, respectively.

This process of adding noise is illustrated in Figure 8, and we select four noise levels (from 20 to 80 dB with an interval of 20 dB) in this article. The newly classification accuracy results among different methods are compared in Figure 9, and it can be observed that the proposed method still has the best performance, which demonstrates the anti-noising capacity.

In addition, to further demonstrate the efficacy of the MANN, we conducted the experiments under Case 7 to develop a new training and testing dataset, that is, the scenarios of Cases 1–6 are repeated once again at Joint B, instead of Joint A. A typical and perpetual problem that current percussion-based methods and

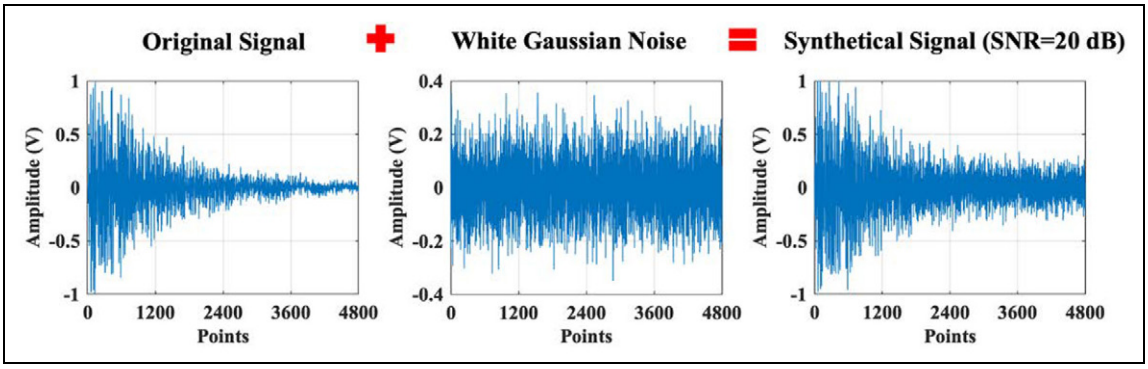

Figure 8. Illustration of the process of adding noise (under the case of SNR = 20 dB).

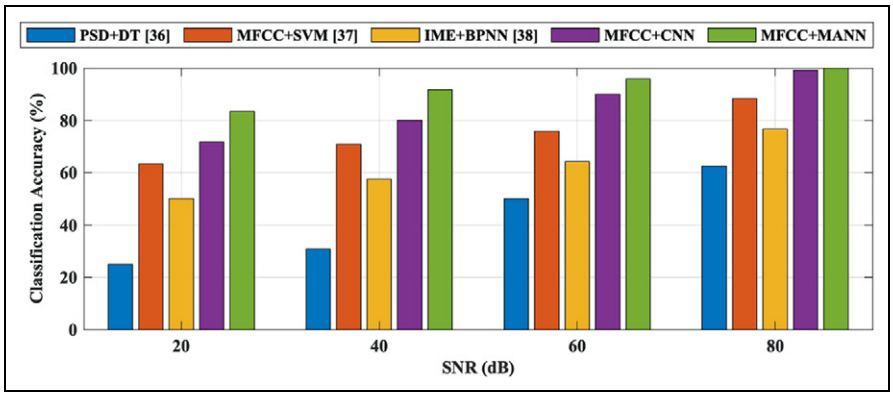

Figure 9. Classification accuracy among different methods.

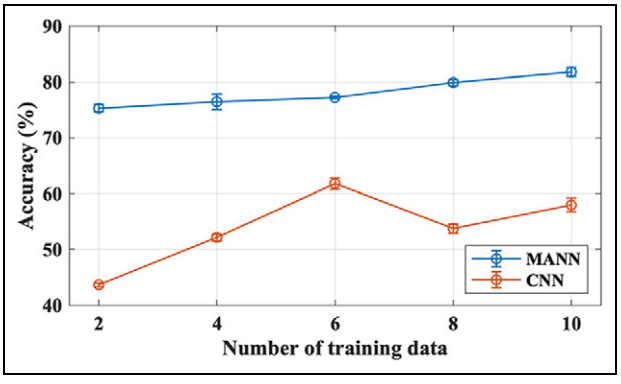

Figure 10. Testing accuracies under Case 7 using the MANN and the CNN.

deep network-based classifiers (e.g. CNN) face is the dramatic deterioration of classification accuracy under a new dataset. On the other hand, a classifier retraining, which is a common solution, is less efficient and costly, since it is difficult to obtain enough training data from all the connections. Therefore, the MANN provides a better choice: we only need a few training data from Joint B to amend the well-trained model (via Joint A) to achieve promising classification accuracy. In other words, no retraining is required, which improves the practicability. For instance, after conducting the new dataset through Case 7, we allow both the MANN and CNN to use only a few new MFCC samples from each class of Joint B for adaptation (here, the number of training set is set to 2, 4, 6, 8, and 10), and Figure 10 compares their classification accuracies of using the new testing set of 20 samples. We can observe that the MANN shows capacity superior to CNN in classification, particularly when only a few new training data are available.

Conclusion

In this article, inspired by the sound-based human diagnostic approach, we propose a novel robotic-assisted percussion method to detect bolt-ball joint looseness in spatial structures. In other words, similar to the exclaim of patients, the percussion-induced sound signals are rich in different features, which can be utilized to detect bolt looseness. Therefore, the MFCC, which is a standard technology for speech recognition, was applied to

extract feature sets from sound signals, and the MANN was used to train the model for classification of different bolt integrity status. The main contribution of this article is the detection of multi-bolt looseness for the first time. Compared to current percussion-based methods, the proposed method can achieve higher accuracy; particularly, it can maintain a good performance when new data are encountered (i.e. when other connections with the same size and configuration are detected). Since this article preliminarily explores and verifies the effectiveness of the robotic-assisted percussion method, we have enough confidence that an automatic inspection system, which provides good guidance for the development of the cyber-physical system, can be developed to detect structural damage. Moreover, in the future work, we will conduct further investigation to determine the sensitivity of the proposed method, such as the minimum looseness degree or the case that the joint is partially loose.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The first author is financially supported by the China Scholarship Council (No. 201706060203), and this research is partially supported by an internal University of Houston grant entitled "Synergistic Research in Intelligent Bolt and Interface Inspection and Monitoring.

ORCID iDs

Furui Wang https://orcid.org/0000-0001-8044-8339
Gangbing Song https://orcid.org/0000-0001-5135-5555

Supplemental material

Supplemental material for this article is available online.

References

1. Han Q, Liu Y and Xu Y. Stiffness characteristics of joints and influence on the stability of single-layer latticed domes. *Thin-wall Struct* 2016; 107: 514–525.
2. Fülöp A and Iványi M. Experimentally analyzed stability and ductility behavior of a space-truss roof system. *Thin-wall Struct* 2004; 42: 309–320.
3. Ghasemi M, Davoodi MR and Mostafavian SA. Tensile stiffness of mero-type connector regarding bolt tightness. *J Appl Sci* 2010; 10: 724–730.
4. Lopez A, Puente I and Serna MA. Direct evaluation of the buckling loads of semi-rigidly jointed single-layer latticed domes under symmetric loading. *Eng Struct* 2017; 29: 101–109.
5. Ebadi M and Davoodi M. Evaluate axial stiffness of the mero connection, under the effect of hardening the screw. *Int J Sci Emerging Tech* 2012; 4: 116–122.
6. Ma H, Fan F, Wen P, et al. Experimental and numerical studies on a single-layer cylindrical reticulated shell with semi-rigid joints. *Thin-wall Struct* 2015; 86: 1–9.
7. Yang X, Lei H and Chen YF. Constant amplitude fatigue test research on M20 high-strength bolts in grid structure with bolt-sphere joints. *Adv Struct Eng* 2017; 20: 1466–1475.
8. Hiyama Y, Takashima H, Iijima T, et al. Buckling behavior of aluminum ball Jointed single layered reticular domes. *Int J Space Struct* 2000; 15: 81–94.
9. Zeng Q, Guo X, Huang Z, et al. Uniaxial compression bearing capacity of bolted ball-cylinder joint. *Eng Struct* 2019; 183: 976–986.
10. Ma H, Ma Y, Yu Z, et al. Experimental and numerical research on gear-bolt joint for free-form grid spatial structures. *Eng Struct* 2017; 148: 522–540.
11. Moreno-Gomez A, Amezquita-Sanchez JP, Valtierra-Rodriguez M, et al. EMD-Shannon entropy-based methodology to detect incipient damages in a truss structure. *Appl Sci* 2018; 8: 2068.
12. An Y, Blachowski B, Zhong Y, et al. Rank-revealing QR decomposition applied to damage localization in truss structures. *Struct Control Health Monit* 2017; 24: e1849.
13. Yu L and Zhu JH. Structural damage prognosis on truss bridges with end connector bolts. *J Eng Mech* 2017; 143: B4016002.
14. Gholizad A and Safari H. Two-dimensional continuous wavelet transform method for multidamage detection of space structures. *J Perform Constr Facil* 2016; 30: 04016064.
15. Li Q and Jing X. Fault diagnosis of bolt loosening in structures with a novel second-order output spectrum–based method. *Struct Health Monit* 2020; 19: 123–141.
16. Zhang T, Biswal S and Wang Y. SHMnet: condition assessment of bolted connection with beyond human-level performance. *Struct Health Monit*. Epub ahead of print 17 October 2019. DOI: 10.1177/1475921719881237.
17. Huo L, Wang F, Li H, et al. A fractal contact theory based model for bolted connection looseness monitoring using piezoelectric transducers. *Smart Mater Struct* 2017; 26: 104010.
18. Wang F, Huo L and Song G. A piezoelectric active sensing method for quantitative monitoring of bolt loosening using energy dissipation caused by tangential damping based on the fractal contact theory. *Smart Mater Struct* 2018; 27: 015023.
19. Wang F, Ho SCM and Song G. Monitoring of early looseness of multi-bolt connection: a new entropy-based active sensing method without saturation. *Smart Mater Struct* 2019; 28: 10LT01.
20. Wang F, Ho SCM, Huo L, et al. A novel fractal contact-electromechanical impedance model for quantitative monitoring of bolted joint looseness. *IEEE Access* 2018; 6: 40212–40220.
21. Huynh T and Kim J. Quantification of temperature effect on impedance monitoring via PZT interface for prestressed tendon anchorage. *Smart Mater Struct* 2017; 26: 125004.

22. Wang F and Song G. Bolt early looseness monitoring using modified vibro-acoustic modulation by time-reversal. *Mech Syst Signal Process* 2019; 130: 349–360.
23. Zhang Z, Liu M, Liao Y, et al. Contact acoustic nonlinearity (CAN)-based continuous monitoring of bolt loosening: hybrid use of high-order harmonics and spectral sidebands. *Mech Syst Signal Process* 2018; 103: 280–294.
24. Wang F, Chen Z and Song G. Monitoring of multi-bolt connection looseness using entropy-based active sensing and genetic algorithm-based least square support vector machine. *Mech Syst Signal Process* 2020; 136: 106507.
25. Fierro GPM and Meo M. IWSHM 2017: Structural health monitoring of the loosening in a multi-bolt structure using linear and modulated nonlinear ultrasound acoustic moments approach. *Struct Health Monit* 2018; 17: 1349–1364.
26. Yang Y, Ng CT and Kotousov A. Bolted joint integrity monitoring with second harmonic generated by guided waves. *Struct Health Monit* 2019; 18: 193–204.
27. Wang F and Song G. Monitoring of multi-bolt connection looseness using a novel vibro-acoustic method. *Nonlinear Dyn* 2020; 100: 243–254.
28. Sun H, Wang T, Liu Q, et al. A novel eddy current array sensing film for quantitatively monitoring hole-edge crack growth in bolted joints. *Smart Mater Struct* 2018; 28: 015018.
29. Xu J, Wang C, Li H, et al. Health monitoring of bolted spherical joint connection based on active sensing technique using piezoceramic transducers. *Sensors* 2018; 18: e1727.
30. Xu J, Dong J, Li H, et al. Looseness monitoring of bolted spherical joint connection using electro-mechanical impedance technique and BP neural networks. *Sensors* 2019; 19: 1906.
31. Ramana L, Choi W and Cha YJ. Fully automated vision-based loosened bolt detection using the Viola-Jones algorithm. *Struct Health Monit* 2019; 18: 422–434.
32. Tong F, Xu XM, Luk BL, et al. Evaluation of tile-wall bonding integrity based on impact acoustics and support vector machine. *Sensor Actuat A-Phys* 2008; 144: 97–104.
33. Delgado-Arredondo PA, Morinigo-Sotelo D, Osornio-Rios RA, et al. Methodology for fault detection in induction motors via sound and vibration signals. *Mech Syst Signal Process* 2017; 83: 568–589.
34. Glowacz A. Fault diagnosis of single-phase induction motor based on acoustic signal. *Mech Syst Signal Process* 2019; 117: 65–80.
35. Wang F, Ho SCM and Song G. Modeling and analysis of an impact-acoustic method for bolt looseness identification. *Mech Syst Signal Process* 2019; 133: 106249.
36. Kong Q, Zhu J, Ho SCM, et al. Tapping and listening: a new approach to bolt looseness monitoring. *Smart Mater Struct* 2018; 27: 07LT02.
37. Zhang Y, Zhao X, Sun X, et al. Bolt loosening detection based on audio classification. *Adv Struct Eng* 2019; 22: 2882–2891.
38. Yuan R, Lv Y, Kong Q, et al. Percussion-based bolt looseness monitoring using intrinsic multiscale entropy analysis and BP neural network. *Smart Mater Struct* 2019; 28: 125001.
39. Lattanzi D and Miller G. Review of robotic infrastructure inspection systems. *J Infrastruct Syst* 2017; 23: 04017004.
40. Myung H, Jung J and Jeon H. Robotic SHM and model-based positioning system for monitoring and construction automation. *Adv Struct Eng* 2012; 15: 943–954.
41. Adams RD, Cawley P, Pye CJ, et al. A vibration technique for non-destructively assessing integrity of structures. *J Mech Eng Sci* 1978; 20: 93–100.
42. Lee H, Pham P, Largman Y, et al. Unsupervised feature learning for audio classification using convolutional deep belief networks. In: *Proceedings of the 22nd international conference on neural information processing system*, vol. 22, Vancouver, BC, Canada, 7–10 December 2009. NY, USA: Curran Associates Inc., pp. 1096–1104.
43. Hau D and Chen K. Exploring hierarchical speech representations using a deep convolutional neural network. In: *Proceedings of the 11th UK workshop computational intelligence*, Manchester, UK, 7–9 September 2011, http://ukci.cs.manchester.ac.uk/files/Proceedings.pdf
44. Abdeljaber O, Sassi S, Avci O, et al. Fault detection and severity identification of ball bearings by online condition monitoring. *IEEE T Ind Electron* 2019; 66: 8136–8147.
45. Liu Z, Wu Z, Li T, et al. GMM and CNN hybrid method for short utterance speaker recognition. *IEEE T Ind Inform* 2018; 14: 3244–3252.
46. Phan H, Hertel L, Maass M, et al. Improved audio scene classification based on label-tree embeddings and convolutional neural networks. *IEEE T Audio, Speech, Lang Process* 2017; 25: 1278–1290.
47. Acar E, Hopfgartner F and Albayrak S. Understanding affective content of music videos through learned representations. In: *International conference on multimedia modeling*, Dublin, 6–10 January 2014, pp. 303–314. New York: Springer.
48. Chowdhury A and Ross A. Extracting sub-glottal and supra-glottal features from MFCC using convolutional neural networks for speaker identification in degraded audio signals. In: *IEEE International joint conference on biometrics (IJCB)*, Denver, CO, 1–4 October 2017, pp. 1–9. New York: IEEE.
49. Jin G, Ye B, Wu Y, et al. Vehicle classification based on seismic signatures using convolutional neural network. *IEEE Geosci Remote S* 2018; 16: 628–632.
50. Abdel-Hamid O, Mohamed AR, Jiang H, et al. Convolutional neural networks for speech recognition. *IEEE Trans Audio, Speech, Lang Process* 2014; 22: 1533–1545.
51. Sehgal A and Kehtarnavaz N. A convolutional neural network smartphone app for real-time voice activity detection. *IEEE Access* 2018; 6: 9017–9026.
52. Noda K, Yamaguchi Y, Nakadai K, et al. Audio-visual speech recognition using deep learning. *Appl Intell* 2015; 42: 722–737.
53. Li TLH, Chan AB and Chun AHW. Automatic musical pattern feature extraction using convolutional neural network. In: *Proceedings of the international multiconference of engineers and computer scientists 2010, IMECS*,

Kowloon, Hong Kong, 17–19 March 2010, pp. 546–550. Kowloon, Hong Kong: City University of Hong Kong.
54. Eghbal-zadeh H, Lehner B, Dorfer M, et al. A hybrid approach with multi-channel I-vectors and convolutional neural networks for acoustic scene classification, 2017, https://arxiv.org/abs/1706.06525
55. Yu Y, Tang S, Raposo F, et al. Deep cross-modal correlation learning for audio and lyrics in music retrieval. *ACM T Multim Comput* 2019; 15: 20.
56. Jiang F, Li H, Zhang Z, et al. An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN. In: *International conference on optical instruments and technology: advanced optical sensors and applications*, vol. 10618, Beijing, China, 28–30 October 2017. Bellingham, WA: SPIE.
57. Huang HM, Chen WK, Liu CH, et al. Singing voice detection based on convolutional neural networks. In: *7th international symposium on next generation electronics (ISNE)*, Taipei, 7–9 May 2018. New York: IEEE.
58. Bahoura M. Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. *Comput Biol Med* 2009; 39: 824–843.
59. Sandipan C, Anindya R and Goutam S. Improved closed set text-independent speaker identification by combining MFCC with evidence from flipped filter banks. *Int J Signal Process* 2007; 4: 114–122.
60. Ai OC, Hariharan M, Yaacob S, et al. Classification of speech dysfluencies with MFCC and LPCC features. *Expert Syst Appl* 2012; 39: 2157–2165.
61. Marcus G. Deep learning: a critical appraisal, 2018, https://arxiv.org/abs/1801.00631
62. Hochreiter S, Younger AS and Conwell PR. Learning to learn using gradient descent. In: *Proceedings of the international conference on artificial neural networks (ICANN)*, Vienna, 21–25 August 2001, pp. 87–94. New York: Springer.
63. Graves A, Wayne G and Danihelka I. Neural Turing Machines, 2014, https://arxiv.org/abs/1410.5401
64. Weston J, Chopra S and Bordes A. Memory networks, 2014, https://arxiv.org/abs/1410.3916
65. Santoro A, Bartunov S, Botvinick M, et al. Meta-learning with memory-augmented neural networks, 2016, pp. 1842–1850, http://proceedings.mlr.press/v48/santoro16.pdf
66. Gopalan R, Li R and Chellappa R. Domain adaptation for object recognition: an unsupervised approach. In: *Proceedings of the international conference on computer vision*, Barcelona, 6–13 November 2011, pp. 999–1006. New York: IEEE.
67. Mobiny A, Moulik S and Nguyen HV. Lung cancer screening using adaptive memory-augmented recurrent networks, 2017, https://arxiv.org/abs/1710.05719
68. Graves A, Wayne G, Reynolds M, et al. Hybrid computing using a neural network with dynamic external memory. *Nature* 2016; 538: 471–476.
69. Yu S, Indurthi S, Back S, et al. A multi-stage memory augmented neural network for machine reading comprehension. In: *Proceedings of the workshop on machine reading for question answering*, 2018, pp. 21–30. Association for Computational Linguistics, https://www.aclweb.org/anthology/W18-2603.pdf
70. Bercea CI, Pauly O, Maier A, et al. SHAMANN: shared memory augmented neural networks. In: *International conference on information processing in medical imaging*, Hong Kong, China, 2–7 June 2019, pp. 830–841.