Structural Health Monitoring

If Structure Can Exclaim: A Novel Robotic-assisted Percussion Method for Spatial Bolt-ball Joint Looseness Detection

Journal:	*Structural Health Monitoring*
Manuscript ID	SHM-20-0025.R1
Manuscript Type:	Machine Learning SI
Date Submitted by the Author:	26-Mar-2020
Complete List of Authors:	Wang, Furui; University of Houston, Mobiny, Aryan; University of Houston Nguyen, Hien; University of Houston Song, gangbing; University of Houston, Department of Mechanical Engineering
Keywords:	Bolt looseness detection, percussion method, structural health monitoring, Mel frequency cepstral coefficient, memory-augmented neural network

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 (MFCC) and the memory-augment neural network (MANN) in this paper. In comparison to current investigations, the main contribution of this paper 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 MANN 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.

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.

Video.mov

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,*]
[1]Department of Mechanical Engineering, University of Houston, Houston, TX 77204, USA
[2]Department of Electrical & Computer Engineering, University of Houston, Houston, TX 77204, USA
[*] Corresponding author: gsong@uh.edu

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 (MFCC) and the memory-augment neural network (MANN) in this paper. In comparison to current investigations, the main contribution of this paper 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 MANN 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

Acronyms

Acronym	Full Name	Acronym	Full Name
ADAM	Adaptive Moment Estimation	ASR	Automatic Speech Recognition
BPNN	Backpropagation Neural Network	CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform	DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform	DT	Decision Tree
EMD	Empirical Mode Decomposition	EMI	Electro-mechanical Impedance
FCM	Fuzzy C-means	GMM	Gaussian Mixture Model
GPU	Graphics Processing Unit	HMM	Hidden Markov Model
IME	Intrinsic Multiscale Entropy	LRUA	Least Recently Used Access
LSTM	Long Short-term Memory	LTE	Label-tree Embedding
MANN	Memory-augment Neural Network	MFCC	Mel-frequency Cepstrum
NTM	Neural Turing Machine	RRQR	Rank-revealing QR
PSD	Power Spectral Density	SHM	Structural Health Monitoring
SMS	Smart Materials and Structures	SVM	Support Vector Machine

1. Introduction

Due to its advantages including beauteous appearance, light-weight, 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 collapses 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.

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, e.g., 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. employed active sensing [29] and EMI methods [30] to detect the looseness of spatial bolt-ball joints, which demonstrated excellent performance. However, instead of multiple loosening, [30] focuses 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 other hand, 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.

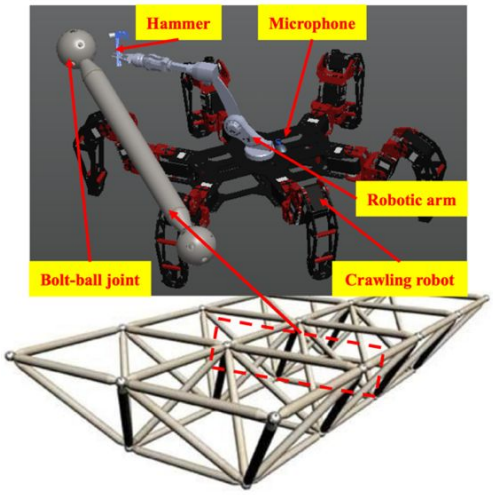

Figure 1 Schematic of robotic-assisted detection of spatial bolt-ball joint looseness

The main contribution of this paper 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 cepstrum (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 to 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 paper is organized as follows. An overview of related literature is presented in Section 2. The theoretical background of the MFCC and the MANN are introduced in Section 3. Section 4 describes the experimental apparatus in detail. The results are discussed in Section 5, and Section 6 concludes this paper. Overall, the proposed method opens up a new frontier for further investigations on the detection of spatial bolt-ball joint loosening.

2. 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. Via this review, we can fully confirm the main contribution of the proposed method, which can be found in the following subsection in detail.

2.1. 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 the 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.

2.2. MFCC features 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 convolutional neural networks (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) 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 require large amounts of data to train the models extensively, and the models need to re-learn 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 paper to obviate the downsides of current methods. More details of MANN are presented in Section 3.2.

3. Methods

The flowchart of the proposed method in this paper is illustrated in Figure 2. Firstly, 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 an LSTM controller and an external memory. More details of the MFCC and the MANN will be introduced in the following subsection.

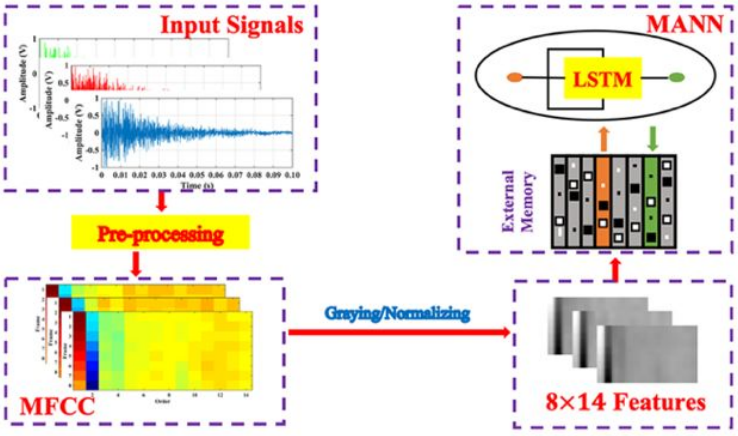

Figure 2 Flowchart of the proposed method

3.1. 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 Mel frequency cepstral coefficient (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 $0\le k\le M_s$, 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,

$$H_i(k)=\begin{cases}0 & k<k_{b_i-1}\\ \dfrac{k-k_{b_i-1}}{k_{b_i}-k_{b_i-1}} & k_{b_i-1}\le k\le k_{b_i}\\ \dfrac{k_{b_i+1}-k}{k_{b_i+1}-k_{b_i}} & k_{b_i}\le k\le k_{b_i+1}\\ 0 & k>k_{b_i+1}\end{cases} \tag{2}$$

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

where k_{b_i} denotes the boundary points of the filter [53]; Q is the total of filters; $f_{\max}$ 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,

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

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

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

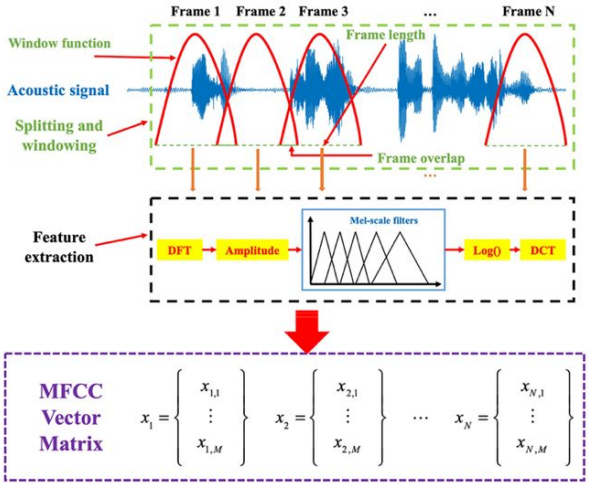

Figure 3 Flowchart of the MFCC feature extraction and computation process

3.2. 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 is 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. developed the Memory-Augmented Neural Networks (MANN) [65], which has a better capacity to accomplish meta-learning [66]. Subsequently, several investigations [67-70] have proven the advantages of MANN.

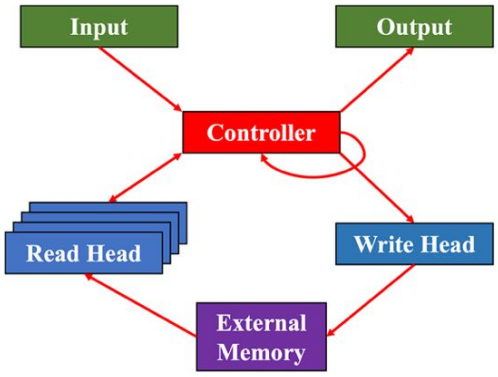

Figure 4 The architecture of the MANN [65]

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; $\mathbf{W}^{xh}(x_t, y_{t-1})$ represents the weight from the input (x_t, y_{t-1}) to the hidden state; $\mathbf{W}^{hh}$ means the weight between two different hidden states; $\mathbf{h}_{t-1}$ is the hidden state under data label y_{t-1}; $\mathbf{b}^h$ is the bias of the hidden state.

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

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

where $\mathbf{h}_t = \sigma(\hat{g}^o) \odot \tanh(\mathbf{c}_t)$; $\mathbf{c}_t = \sigma(\hat{g}^f) \odot \mathbf{c}_{t-1} + \sigma(\hat{g}^i) \odot \tanh(\hat{\mathbf{u}})$ denotes the cell state; $\mathbf{r}_t$ is the read vector according to the external memory $\mathbf{M}_t$; $\sigma()$ represents 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 $K(k_t, \mathbf{M}_t(i)) = \mathbf{k}_t \cdot \mathbf{M}_t(i) / \|\mathbf{k}_t\| \|\mathbf{M}_t(i)\|$ 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,

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

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

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

4. 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: 0.35m length, 0.35m width, and 2.1m 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. 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 0.2m) to reduce the effect of airflow, and its sampling time and sampling rate are set to 0.1 s and 48 kHz, respectively.

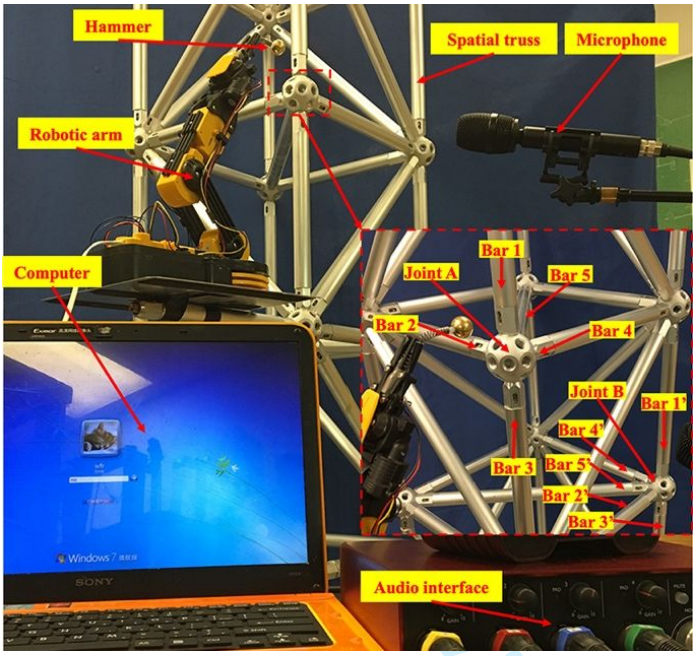

Figure 5 Experimental setup

In this paper, 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 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 other hand, 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.

Table 1 Details of different experimental scenarios

Case	Joint	Looseness	Tightened (20 Nm)
1	A	N/A	Bar 1, 2, 3, 4, and 5
2		Bar 1 (10 Nm)	Bar 2, 3, 4, and 5
3		Bar 1 (0 Nm)	Bar 2, 3, 4, and 5
4		Bar 1 and 2 (0 Nm)	Bar 3, 4, and 5
5		Bar 1, 2 and 3 (0 Nm)	Bar 4 and 5
6		Bar 1, 2, 3 and 4 (0 Nm)	Bar 5
7	B	Corresponding to cases 1-6	

5. Results and Discussion

After down-mixing the percussion-induced sound signals to the mono channel (taking mean values from the left and right channel), we depicted samples under Case 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 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).

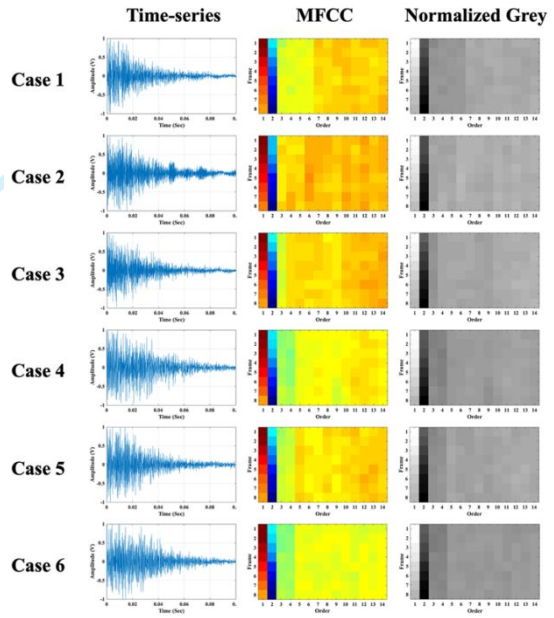

Figure 6 Samples of sound signal and MFCC features under different cases

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 paper, 128 memory slots with a size of 40 were found to be the optimal solution, and the size of the LSTM controller was 200. Additionally, 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 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 can find that the MANN achieves the highest testing accuracy after 25, 000 episodes with 100 % through the tenth 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 were 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.

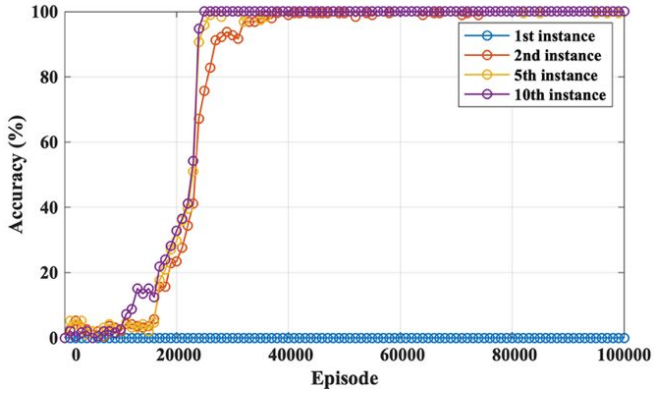

Figure 7 Testing accuracies under different cases (1-6) using the MANN

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**

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

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

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

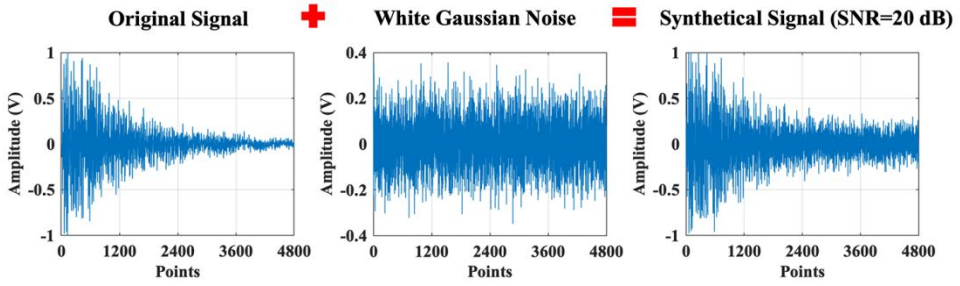

Figure 8 Illustration of the adding noise process (under the case of SNR= 20dB)

This process of adding noise is illustrated in Figure 8, and we select four noise levels (from 20 dB to 80 dB with an interval of 20 dB) in this paper. 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.

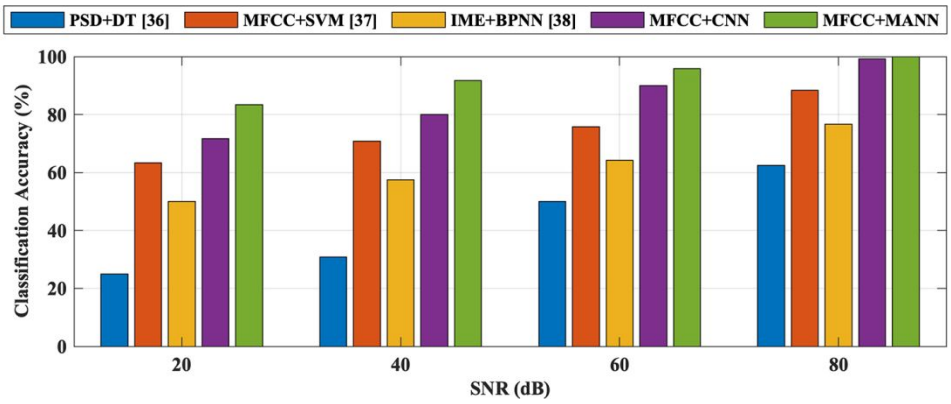

Figure 9 Classification accuracy among different methods

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, i.e., 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 deep network-based classifiers (e.g., CNN) face is the dramatical 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 is available.

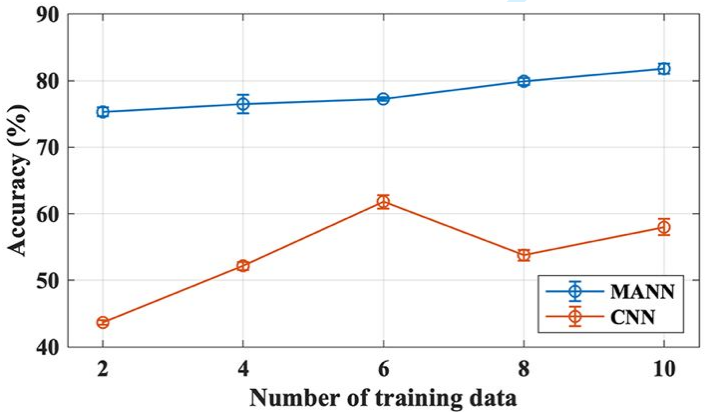

Figure 10 Testing accuracies under case 7 using the MANN and the CNN

6. Conclusions

In this paper, 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 paper 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 is encountered (i.e., when other connections with the same size and configuration are detected). Since this paper 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.

Acknowledgement

This research was partially supported by the China Scholarship Council (No. 201706060203), their financial support is appreciated.

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 signle-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, Oct.

[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 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 Modulation Method. Nonlinear Dyn 2020; DOI: 10.1007/s11071-020-05508-7.
[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 Proc Adv Neural Inf Process Syst* 2009; 22: 1096–1104.
[43] Hau D and Chen K. Exploring hierarchical speech representations using a deep convolutional neural network. *In Proc 11th UK Workshop Comput Intell* 2011.
[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 Trans 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 Proc 20th Int Conf MultiMedia Model* 2014; 303–314.
[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 Proc IJCB* 2017; 1-9.
[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 Proc Int Conf Data Mining Appl* 2010; 546-550.
[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. arXiv preprint arXiv:1706.06525, 2017.
[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 Proc SPIE*, 2018; 10618: 1061804.

[57] Huang HM, Chen WK, Liu CH, et al. Singing voice detection based on convolutional neural networks. *In 2018 7th IEEE ISNE* 2018.

[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. arXiv preprint arXiv:1801.00631, 2018.

[62] Hochreiter S, Younger AS and Conwell PR. Learning to learn using gradient descent. *In Proc Int Conf Artificial Neural Networks (ICANN)* 2001; 87–94.

[63] Graves A, Wayne G and Danihelka I. Neural Turing Machines. arXiv preprint arXiv:1410.5401, 2014.

[64] Weston J, Chopra S and Bordes A. Memory networks. arXiv preprint arXiv:1410.3916, 2014.

[65] Santoro A, Bartunov S, Botvinick M, et al. Meta-learning with Memory-Augmented Neural Networks. *In Proc 33rd Int Conf Machine Learning* 2016; 1842–1850.

[66] Gopalan R, Li R and Chellappa R. Domain adaptation for object recognition: An unsupervised approach. *In Proc IEEE Int Conf Comput Vis* 2012; 999–1006.

[67] Mobiny A, Moulik S and Nguyen HV. Lung Cancer Screening Using Adaptive Memory-Augmented Recurrent Networks. arXiv preprint arXiv:1710.05719, 2017.

[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 Proc of the Workshop on Machine Reading for Question Answering at Association for Computational Linguistics* 2018; 21-30.

[70] Bercea CI, Pauly O, Maier A, et al. SHAMANN: Shared Memory Augmented Neural Networks. *In International Conference on Information Processing in Medical Imaging* 2019; 830-841.