Smart Structures and Systems, *Vol. 29, No. 4 (2022) 625-640*
https://doi.org/10.12989/sss.2022.29.4.625

Multiple damage detection of maglev rail joints using time-frequency spectrogram and convolutional neural network

Su-Mei Wang [1,2], Gao-Feng Jiang [1,2], Yi-Qing Ni*[1,2], Yang Lu [1,2], Guo-Bin Lin [3,4], Hong-Liang Pan [4], Jun-Qi Xu [3,4] and Shuo Hao [1,2]

[1] *Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong S.A.R.*
[2] *National Rail Transit Electrification and Automation Engineering Technology Research Center (Hong Kong Branch), Hung Hom, Kowloon, Hong Kong S.A.R.*
[3] *Shanghai Key Laboratory of Rail Infrastructure Durability and System Safety, Tongji University, Shanghai, China*
[4] *Maglev Transportation Engineering R&D Center, Tongji University, Shanghai, China*

(*Received June 27, 2021, Revised December 30, 2021, Accepted January 3, 2022*)

Abstract. Maglev rail joints are vital components serving as connections between the adjacent F-type rail sections in maglev guideway. Damage to maglev rail joints such as bolt looseness may result in rough suspension gap fluctuation, failure of suspension control, and even sudden clash between the electromagnets and F-type rail. The condition monitoring of maglev rail joints is therefore highly desirable to maintain safe operation of maglev. In this connection, an online damage detection approach based on three-dimensional (3D) convolutional neural network (CNN) and time-frequency characterization is developed for simultaneous detection of multiple damage of maglev rail joints in this paper. The training and testing data used for condition evaluation of maglev rail joints consist of two months of acceleration recordings, which were acquired in-situ from different rail joints by an integrated online monitoring system during a maglev train running on a test line. Short-time Fourier transform (STFT) method is applied to transform the raw monitoring data into time-frequency spectrograms (TFS). Three CNN architectures, i.e., small-sized CNN (S-CNN), middle-sized CNN (M-CNN), and large-sized CNN (L-CNN), are configured for trial calculation and the M-CNN model with excellent prediction accuracy and high computational efficiency is finally optioned for multiple damage detection of maglev rail joints. Results show that the rail joints in three different conditions (bolt-looseness-caused rail step, misalignment-caused lateral dislocation, and normal condition) are successfully identified by the proposed approach, even when using data collected from rail joints from which no data were used in the CNN training. The capability of the proposed method is further examined by using the data collected after the loosed bolts have been replaced. In addition, by comparison with the results of CNN using frequency spectrum and traditional neural network using TFS, the proposed TFS-CNN framework is proven more accurate and robust for multiple damage detection of maglev rail joints.

Keywords: 3D convolutional neural network; damage detection; maglev rail joints; time-frequency spectrogram

1. Introduction

Maglev is a new mode of transportation using non-contact suspension technologies (Ohtsuka and Iguchi 1982, Huang *et al*. 2018). It is essentially divided into electro-dynamic suspension (EDS) (Sato *et al*. 1985, Fujie 1989, Masada 1993) and electromagnetic suspension (EMS) (Sung *et al*. 2006, Goodall 2008, Kim *et al*. 2015, Zhang *et al*. 2017, Li *et al*. 2018) in accordance with the suspension strategy. As the EMS system for medium-low running speed is less expensive than the EDS system and has good ride quality, rapid progress in EMS maglev has been made in China, Japan, South Korea, and other countries. For the EMS system, the functions of suspension and guidance are realized by the electromagnet force between the electro-magnets and guideway (Goodall 2008). To generate the electromagnet force between the on-board electromagnets and guideway, the guideway rail is usually designed to be "F" type (Zhang *et al*. 2017). To ensure comfort and stability of the maglev train, the range of suspension gap fluctuation between the electromagnets and F-type rail is confined within 2 to 3 mm. Such a small fluctuation invokes a high demand for the condition of the F-type rail. In practice, it is observed that large gap disturbance usually occurs at maglev rail joints where two F-type rails are connected at their ends, to prevent movement of the rail horizontally and vertically (Li *et al*. 2018). In addition to bridging adjacent F-type rails, the maglev rail joints also allow for longitudinal movement due to temperature-induced expansion and contraction (Dangre 2019). When a maglev train runs across a rail joint, gap fluctuation at the rear suspension point is usually larger than that at the front suspension point (Zhou *et al*. 2017). This is brought by the so-called rail step, which is a structural characteristic of maglev rail joints and is the main excitation sources of rail irregularity (Alberts *et al*. 2011, Li *et al*. 2018). Another structural characteristic of maglev rail joints is the lateral rail dislocation which leads to an abrupt decrease in the electromagnet force between the electromagnets and F-type

*Corresponding author, Chair Professor,
E-mail: ceyqni@polyu.edu.hk

ISSN: 1738-1584 (Print), 1738-1991 (Online)

rail. The rail step and lateral dislocation at maglev rail joints are mainly caused by the installation inaccuracy, temperature changes, bolt looseness, and ground settlement (Li *et al*. 2020b). When the rail step or lateral dislocation exceeds a certain range, the maglev train vibrates severely, and malfunction of the suspension system may occur. Large impacts caused by repeated gap fluctuations make the maglev rail joints the weakest part in the guideway. Also, the large impacts and dynamic suspension forces acting on the rail joints aggravate structural deterioration (Kabo *et al*. 2006, Oregui *et al*. 2017), and thus bolts get loose, rail ends become battered, and cracks develop in the F-type rail (Oregui *et al*. 2017). Such damage may lead to large vibrations of maglev systems and even clashes between the electromagnets and F-type rail, trimming down the ride comfort of maglev trains. However, invisible damage (e.g., bolt looseness) cannot be detected by the commonly practiced visual inspection by persons, which is labor intensive, unreliable, intrusive, and unsafe (Oregui *et al*. 2017). The condition monitoring of maglev rail joints based on advanced sensing technologies is therefore highly desired for online damage detection and to maintain the safe operation of maglev systems.

With the advance of digital camera technology, vision-based techniques have been proposed for damage detection with regard to structural joints. The vision-based approaches are to feed images/videos into image processing algorithms, which then extract abnormal features indicative of damage or changes relative to a healthy baseline (Dung *et al*. 2019, Pham *et al*. 2020a). With the rapid development in deep learning, various algorithms have emerged with the ability to automatically classify and recognize structural features with satisfactory accuracy (Ye *et al*. 2019). Park *et al*. (2015) developed an image processing technique based on the Hough Transform for looseness detection of bolted joints in steel bridges, which is important for maintaining the structural performance of steel bridges. Then, Zhang *et al*. (2020) improved the technique developed by Park *et al*. (2015) and used region-based convolutional neural network (RCNN) to enhance the accuracy of damage prediction in regard to bolted joints. Cha *et al*. (2017) developed a visual inspection method based on faster region-based convolutional neural network (Faster RCNN) to provide quasi-real-time simultaneous detection of five types of structural damage. As regards structural joints in railway engineering, Marino *et al*. (2007) established a multilayer perception neural classifier to detect the absence of rail fastening bolts by using images from a digital line-scan camera. Gibert *et al*. (2017) applied deep CNN and multi-task learning methods to inspect fasteners and ballast in railway tracks. Chen *et al*. (2018) adopted a deep CNN-based model to detect the defects of fastening joints on rail catenary support devices. All these structural joint damage detection methods rely on images taken from detected objects. However, several shortcomings are found in feature extraction directly from captured photographs. For instance, the image-based detection methods cannot perceive inner structural damage or damage in visually inaccessible locations (Teng *et al*. 2020). The accuracy of detection also varies with the levels and changes of illumination (Li *et al*. 2020a) and angles of camera (Zhao *et al*. 2019). Vision-based structural joint damage detection could be supplemented with information about the structural response features.

Recently, numerous damage detection methods adopting deep learning and structural response features have been proposed. These methods can be divided into parametric and non-parametric methods. The former is model-based while the latter is signal-based (Abdeljaber *et al*. 2017). The parametric methods employ inherent structural properties such as modal parameters for damage detection. For instance, to detect the location of damaged rods in a steel truss model, Zhong *et al*. (2020) used the first-order modal parameters, including mode shapes and mode curvatures, as input to train a CNN model. The non-parametric methods use structural static/dynamic response obtained numerically or experimentally. Lin *et al*. (2017) established a damage detection model in terms of one-dimensional (1D) CNN using the calculated accelerations from the numerical model of a steel beam. Teng *et al*. (2020) detected the damage of a steel frame using a CNN model trained by the numerical results of a finite element model and verified by experimental data. Compared with acceleration responses in the time domain, their frequency-domain counterparts contain more useful information, especially for non-stationary signals (Verstraete *et al*. 2017). Duan *et al*. (2019) used Fourier amplitude spectra of acceleration responses to train a CNN model, and the trained model was proven robust for detecting damage in an arch bridge. Short-time Fourier transform (STFT) is an effective method to articulate both time-domain and frequency-domain information from vibration signals. Pham *et al*. (2020b) applied VGG-16 CNN to diagnose bearing faults and their degradation degree under variable shaft speeds, using spectrograms generated by STFT. Zheng *et al*. (2020) extracted time and frequency features of ground motions from STFT to assess the destructive power of earthquakes to structures by formulating backpropagation neural network (BPNN) and CNN.

To the best of the authors' knowledge, there is a paucity of research addressing damage detection of maglev rail joints. In this regard, the present study aims to develop a CNN-based model for simultaneously identifying multiple damage in maglev rail joints. The acceleration signals acquired from an online monitoring system are transformed to the time-frequency domain by STFT. The time-frequency spectrograms (TFSs) are obtained to train a three-dimensional (3D) CNN model. The trained CNN-based damage detection model is then implemented for both condition assessment and damage detection of maglev rail joints. This paper is structured as follows: Section 2 introduces an overview of maglev rail joints. Section 3 presents the in-situ experiments with an online monitoring system. Section 4 describes the formulation of the TFS-CNN model and the CNN-based multiple damage detection method. The experimental results are presented in Section 5. Section 6 concludes the study.

2. Overview of maglev rail joints

In general, the maglev guideway consists of multi-span bridge(s), steel sleepers, and multi-segment F-type rail. The F-type rail is specially designed to maintain the suspension requirement of maglev trains and is composed of linear motor induction plate, suspension clearance detection surface, internal magnetic pole surface, external magnetic pole surface, and running surface (Li *et al*. 2015), as shown in Fig. 1. The F-type rail, as a portion of maglev guideway, is prone to deformation/deflection due to various causes such as temperature changes, settlement of foundations, and force action. The temperature changes play the most influential role among the various causes (Yau 2009). The growing temperature difference of maglev guideway will bring about great temperature gradient and cause a significant increase of vertical or transverse deformation (Wu *et al*. 2019, Zhang and Huang 2018). Hence, a specific gap between the adjacent F-type rails is considered in the design to allow for expansion of the guideway under such conditions. The gap between two adjacent F-type rails is known as rail seam. On the existing medium-low speed maglev lines, three kinds of maglev rail joints: JI-type, JII-type, and JIII-type, are used to connect the rail seam, as shown in Fig. 2. The JI-type and JII-type rail joints connect the single rail seam and double rail seams on the straight guideway. The JIII-type rail joint is seen on the curved guideway. The JI-type rail joint, studied in this paper, is the most common rail joint in maglev lines and consists of wedge block, cover plate, and bolts.

The main geometrical parameters of JI-type rail joints include the rail step height, lateral dislocation length, and rail seam width, as shown in Fig. 3. The rail step height is defined as the height deviation of the adjacent ends of two F-type rails, the length of lateral dislocation is the lateral deviation of the adjacent ends of two F-type rails, while the rail seam width refers to the gap width between two adjacent F-type rails in the longitudinal direction. The rail seam can cause instability of maglev train. Hence, to ensure stability in operation, the threshold method (Sung *et al*. 2006, Li *et al*. 2018) has been adopted in the control system, in which the minimum gap signal among the three gap sensor probes on the magnet is picked up when the gap sensors go through a maglev rail joint (refer to Fig. 4). Due to the redundant configuration of three suspension gap sensor probes and the fusion strategy of gap signals, the influence of the rail joint can be eliminated in a certain range. However, it is important to note that the rail joint, as the weakest portion of the guideway, is highly vulnerable to damage. If the geometrical parameters of maglev rail joints are changed as a result of rail joint damage, the threshold method will lead to a discontinuous suspension gap at the rail joint (Sung *et al*. 2006). The discontinuity of the suspension gap has the greatest impact on the vibration of maglev rail joints, and this in turn, aggravates the damage of maglev rail joints. The most common damage of JI-type rail joints, which was observed in commercial and testing maglev lines, is the bolt looseness and installation error, which can cause big height of rail step or large length of lateral dislocation. In view of this, a multiple damage detection method is proposed in this paper to detect the operation condition and potential damage of JI-type rail joints.

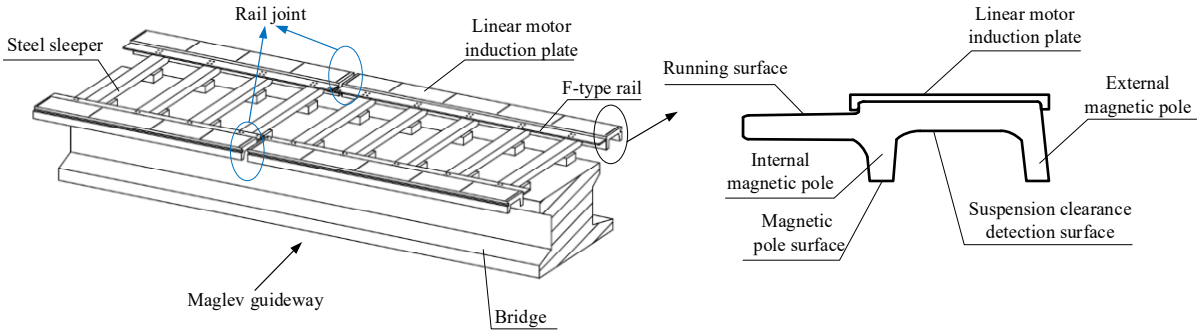


Fig. 1 Configuration of maglev guideway

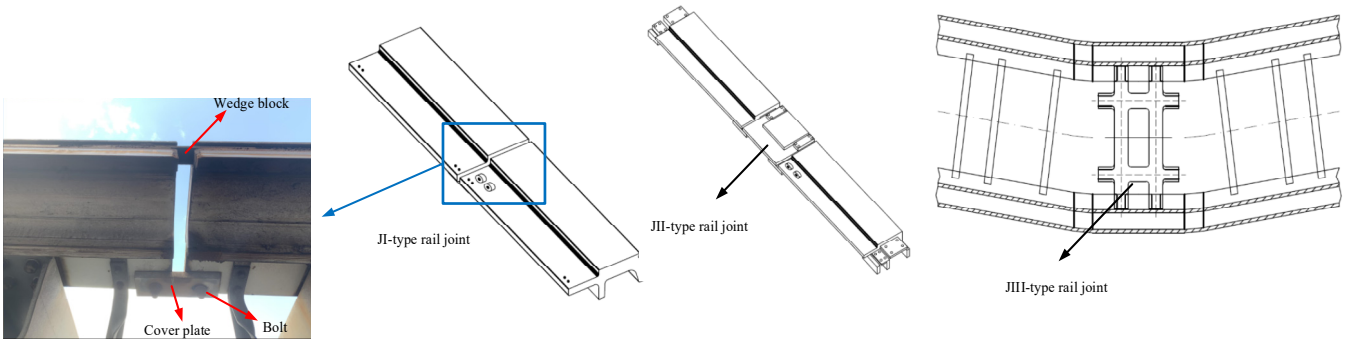


Fig. 2 Three types of maglev rail joints

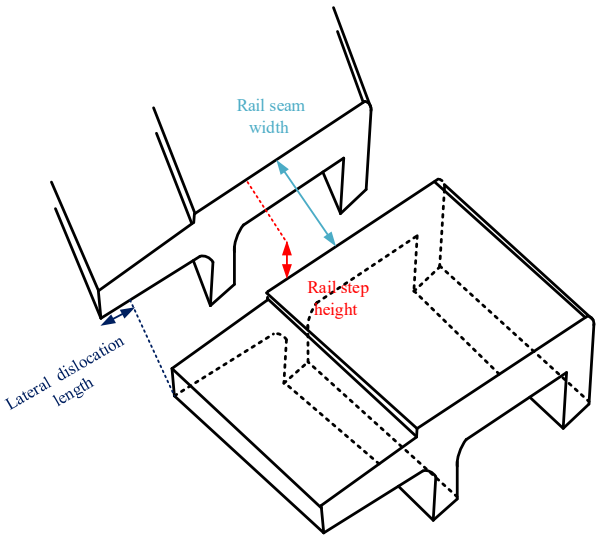


Fig. 3 Main geometrical parameters of JI-type rail joint

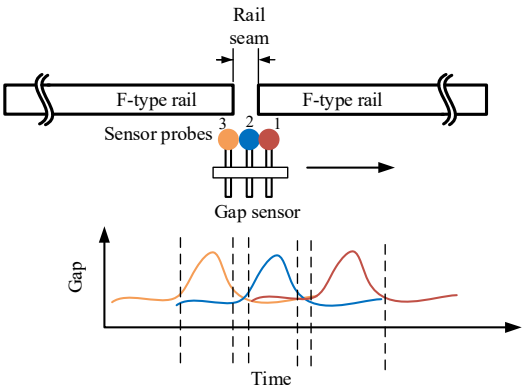


Fig. 4 A gap sensor with three sensor probes going through rail joint

3. Field experiment with online monitoring system

3.1 Monitoring system

An online monitoring system has been developed by the authors for condition monitoring of rail joints. As shown in Fig. 5, the online monitoring system comprises a set of accelerometers, a multiple-channel data acquisition unit, and a high-performance computer server equipped with data processing modules. Because of the electromagnet property of the maglev system, the deployed sensors, signal cables, and data acquisition unit are insulated to immunize them against electromagnetic interference (EMI). Hence, the piezoelectric accelerometers with anti-EMI capability are used. To ensure sufficient signal acquisition resolution to help detect high-frequency ingredients generated by damage, the data acquisition unit is set to collect data at a sampling rate of 5000 Hz. During the maglev train passage, the data acquisition unit is triggered to sample and store data automatically. To facilitate multiple damage detection at different maglev rail joints, a high-performance server with 8-cores, 16-threadings, and 64-Gb memory is adopted.

The system was deployed at a maglev test line in China

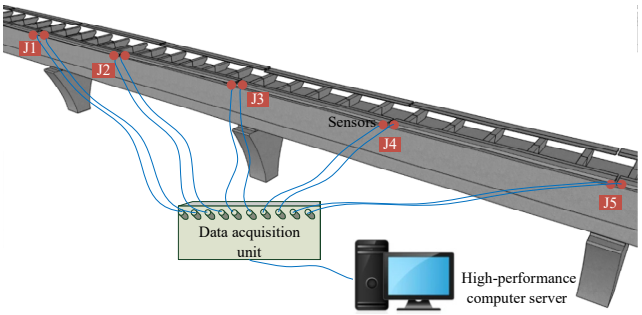


Fig. 5 The devised online monitoring system for maglev rail joints

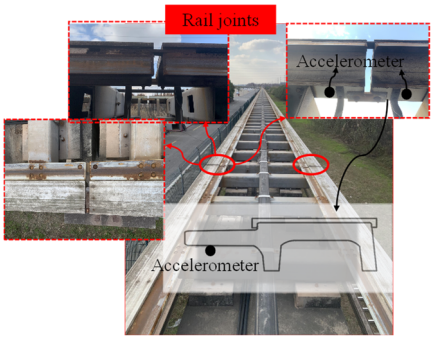


Fig. 6 Deployment of accelerometers near maglev rail joints

and was in service for one year collecting acceleration data intermittently. The total length of the maglev test line is about 1.7 km which consists of multi-span simply supported bridges. The JI-type rail joint is adopted to connect each F-type rail on the bridges. According to reports from the rail operator, there was one rail joint with a lateral dislocation of about 2 mm because of the installation error, and thus the bolts on the rail joint to connect two adjacent rails were most vulnerable to loosen due to the constant impacts from maglev trains. In view of this, as shown in Fig. 6, two accelerometers were mounted on the adjacent ends of two F-type rails and deployed at the cantilevered side (inward part) of the cross-section of the F-type rail. Five rail joints named J1 to J5 were selected to be monitored, covering a monitoring range of about 80 m. All the rail joints can be regarded to work under the same weather condition. The rail joint J2 is the one with lateral dislocation; and the others were in good condition at the time when the system was installed. Additionally, since the vertical electromagnet force is the primary component of the electromagnet force and temperature barely causes a change in acceleration, the sensors were deployed on the F-type rail only for the measurement of accelerations in the vertical direction. The trial running speed of maglev trains moving on the test line ranges from 10 km/h to 80 km/h. To adequately record the rail response, the data acquisition unit took data recording for 30 s for each trial run. Each record is treated as one sample to be used in training or testing of the TFS-CNN model.

The deployed online monitoring system along with the TFS-CNN algorithm aims to provide a viable alternative to automatically inspecting the condition of rail joints. Even

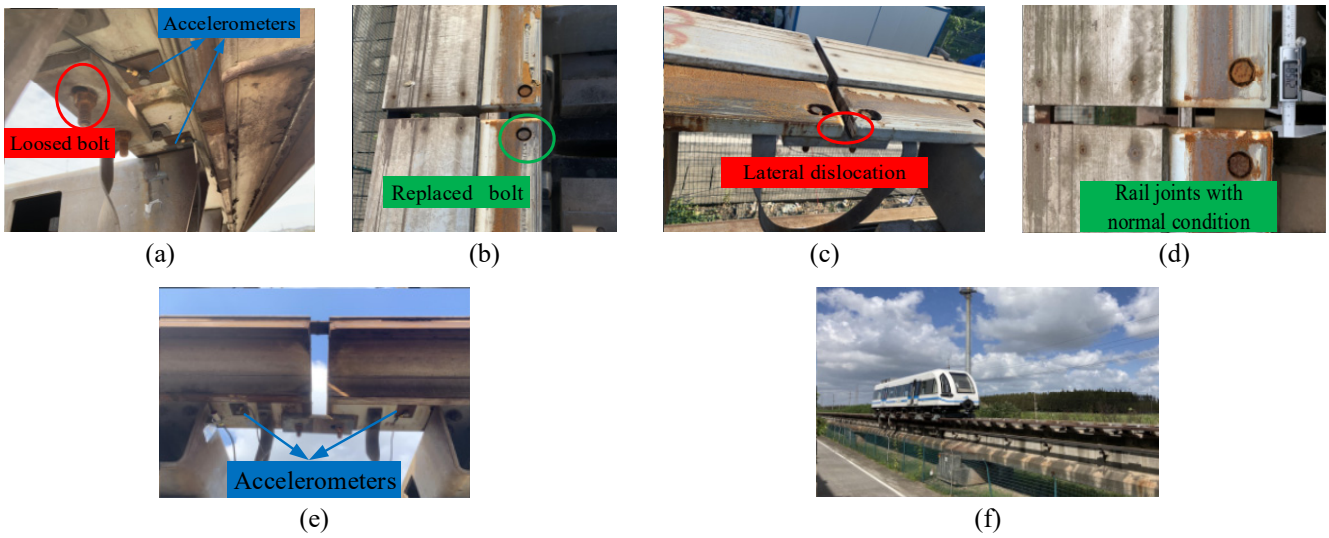
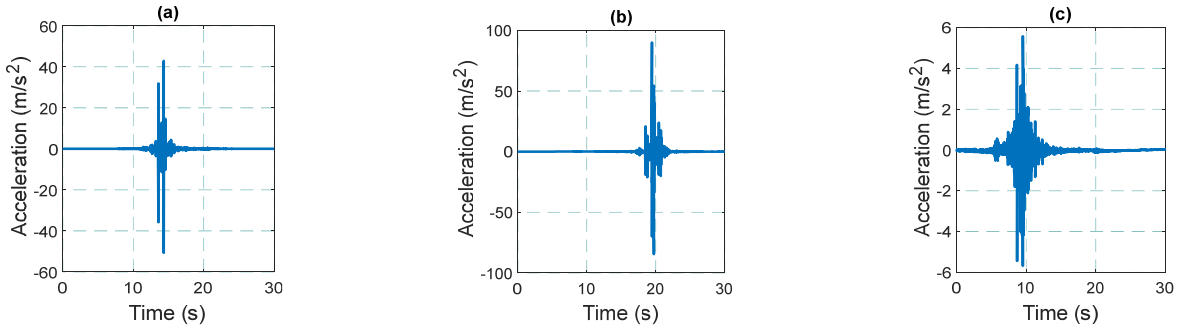


Fig. 7 Rail joints on the maglev test line: (a) rail joint with loosed bolt; (b) rail joint after the loosed bolt replaced; (c) rail joint with lateral dislocation; (d) rail joint with normal condition; (e) accelerometers deployed on a rail joint; (f) a maglev train traveling on the maglev test line



(a) Rail step (b) Lateral dislocation (c) Normal condition

Fig. 8 Accelerations collected on rail joints under different conditions

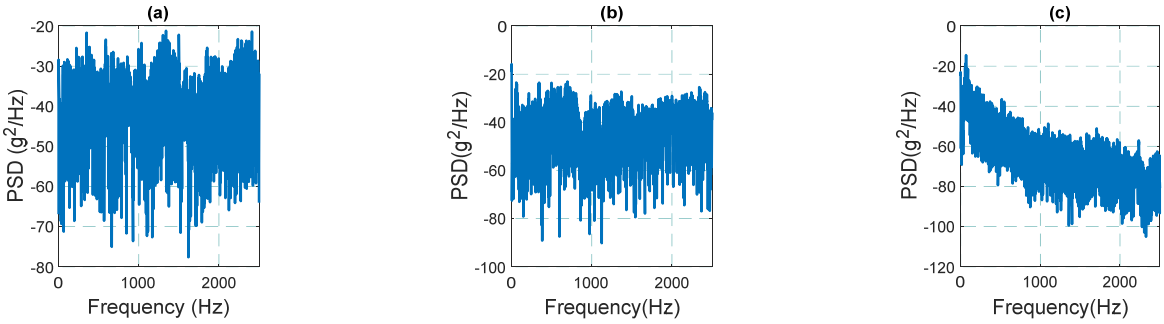
though the developed damage detection model is learned from the data collected from the maglev test line, the method is extendable for multiple damage detection of maglev rail joints in any commercial maglev lines.

3.2 Data description

In early September 2020, a large rail step caused by the looseness of bolts was found on the rail joints J1 and J4 by manual inspection (Fig. 7(a)). Then, the loosed bolts were replaced in mid-October 2020 (Fig. 7(b)). During this period, the conditions of the rail joints J2 (Fig. 7(c)), J3 (Fig. 7(d)) and J5 did not change. The temperature during September (with an average temperature of 27°C) and October (with an average temperature of 23°C) when the data were collected was relatively steady. Therefore, the data recorded from September to October 2020, covering three states of maglev rail joints, were used in this study. The database is divided into three classes in accordance with the three states of maglev rail joints: rail step, lateral dislocation, and normal condition. The data collected from the sensors on the rail joints J1, J2 and J3 in September 2020 are used for model training, and the formulated model is tested and verified by using other data in the database. Fig. 8 shows the acceleration samples of rail joints under three conditions grabbed from the dataset, while the power spectral densities (PSDs) of these signals are illustrated in Fig. 9, concentrating on the frequency range less than 2500 Hz. As can be seen in Fig. 9, both the acceleration amplitude and frequency range of a rail joint with damage is much higher than that of a rail joint under normal condition. It should be noted that Fig. 9 shows only samples grabbed from the database for demonstration and we do not expect to directly observe the different damage states. The procedures for training and testing the CNN model based on these acceleration data will be detailed in Section 4.

4. CNN-based multiple damage detection model

Convolutional neural network (CNN) is well known as a feature extractor for image recognition and classification. With the development of big data analytics and graphic processing units (GPU), CNNs have been widely applied in various fields. In the field of civil and infrastructure health monitoring, CNNs have been applied for road pothole



(a) Rail step (b) Lateral dislocation (c) Normal condition

Fig. 9 PSDs of accelerations collected on rail joints under different conditions

detection (Luo *et al*. 2019), pipeline inspection (Feng *et al*. 2017), crack recognition of steel bridges (Dung *et al*. 2019) and concrete structures (Kim *et al*. 2020, Deng *et al*. 2020), detection of rail surface defects (Faghih-Roohi *et al*. 2016) and slab track (Ma *et al*. 2020). This paper is devoted to developing a CNN-based model which enables the identification of multiple damage in maglev rail joints.

4.1 Input of CNN: time-frequency spectrogram of acceleration data

Before training the CNN model, it is critical to make pre-processing of the collected acceleration data. The measured time-history acceleration data from the monitoring system represent the dynamic responses of maglev rail joints in the time domain. However, the features about structural damage are reflected not only in the time domain but also in the frequency domain. Compared with the acceleration amplitudes in the time domain, the frequency-domain information is more reflective of structural damage. Since the vibration frequency of a maglev rail joint may vary with the running speed and weight of the maglev train, it is necessary to extract the instant features of a maglev rail joint in time and frequency domains simultaneously. Here, the time-frequency spectrogram (TFS) is taken into account, which explicitly describes the features of a signal in both time and frequency domains. To convert the raw time-history signals into TFSs, the short-time Fourier transform (STFT) is conducted that yields

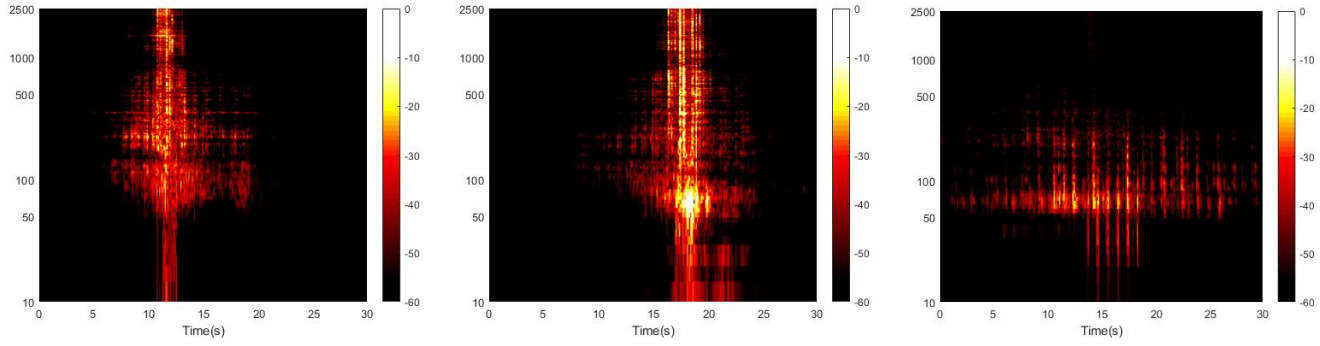
$$X_{STFT}(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]\, e^{-j\omega n} \tag{1}$$

where w [·] is a window function, $x[n]w[n-m]$ is a short-time section of time-history signal $x[n]$ at time m. The time interval for the generation of TFSs is constant to 30 seconds in the present study. The moving Hamming window with the size of 2048 is used to reduce the leakage of the TFSs and alleviate the effect of random noises. After doing so, the PSD of the signal is obtained as

$$X_{PSD}(m,\omega) = 2 \times \frac{|X_{STFT}(m,\omega)|^2}{F_s \sum_{n=1}^{L} |w[n]|^2} \tag{2}$$

where F_s is the sampling frequency and L is the number of discrete Fourier transform (DFT) points, equal to 5000 Hz and 2048, respectively. To better reflect the features in the frequency domain, the PSD is evaluated with the unit of decibel. Note that all TFSs are displayed in RGB image decomposition as the PSD values are expressed as color image.

Fig. 10 shows the TFSs converted from the raw acceleration data of three maglev rail joints under different conditions, i.e., rail step caused by bolt looseness, rail lateral dislocation caused by installation error, and normal condition. It is observed that the vibration frequency of the normal rail joint ranges from 50 Hz to 100 Hz, while higher-frequency components (higher than 100 Hz) appear when the rail joints are under damage condition. According



(a) Rail joint with rail step (b) Rail joint with lateral dislocation (c) Rail joint with normal condition

Fig. 10 The TFSs of three maglev rail joints

to the color bar set in Fig. 10, with the increase of the decibel levels, the color of TFSs becomes bright. Compared to the rail joint with rail step, the levels of PSD for the rail joint with lateral dislocation are much higher between 50 Hz and 100 Hz, basically greater than -20 dB thus generating a large bright area in the TFS. Since the time-frequency features among the three maglev rail joints are discriminatively shown in TFS, it can be concluded that TFS is appropriate as input to formulate a CNN-based model for multiple damage detection.

4.2 Construction of CNN-based model

The architecture of the CNN-based model adopted in this study is schematically illustrated in Fig. 11. It consists of two convolution layers, two max-pooling layers, and two fully connected (FC) layers. The convolution layers and max-pooling layers are set alternately, and responsible for feature extraction, followed by the FC layers for classification. Meanwhile, some improvement strategies are applied to the model classification with the objective of enhancing the performance on multiple damage classification. One of them is data augmentation, which is applied due to the limited initial input TFSs. This strategy produces random geometrical transformations of TFSs, such as rescale, rotate, width and height shift, shear, zoom, and horizontal flip, to expand the dataset of TFSs and reduce the overfitting risk caused by a small-scale dataset. Moreover, some other improvement strategies are joined, including nonlinear activation function, dropout, and batch normalization. More details about the layers of the CNN model and improvement strategies are provided in the following.

4.2.1 Convolution layer

A convolution layer is composed of input and output feature maps, convolution kernel, activation function, and bias. The convolution layer aims to extract various time-frequency features through the convolution process of the input feature map and convolution kernel. The convolution process is defined as

$$E(i,j)=\sum_{m=0}^{M}\sum_{n=0}^{N} I(i+m,j+n)*W(m+1,n+1) \quad (3)$$

where I is the input matrix, W is the convolution kernel, and M and N are the width and height of the convolution kernel W. As shown in Fig. 12, the convolution kernel slides across the input matrix at a fixed stride, where each element in the convolution kernel multiplies the corresponding element of each sub-matrix of input, which is summed to get an element in an output matrix. The output matrix E is formed by repeating the convolution process until all elements of the input matrix are addressed.

After the convolution process, nonlinear activation function and bias are added to the output matrix E to boost the learning of the model. The sigmoid function and Rectified Linear Unit (ReLU) are most commonly used

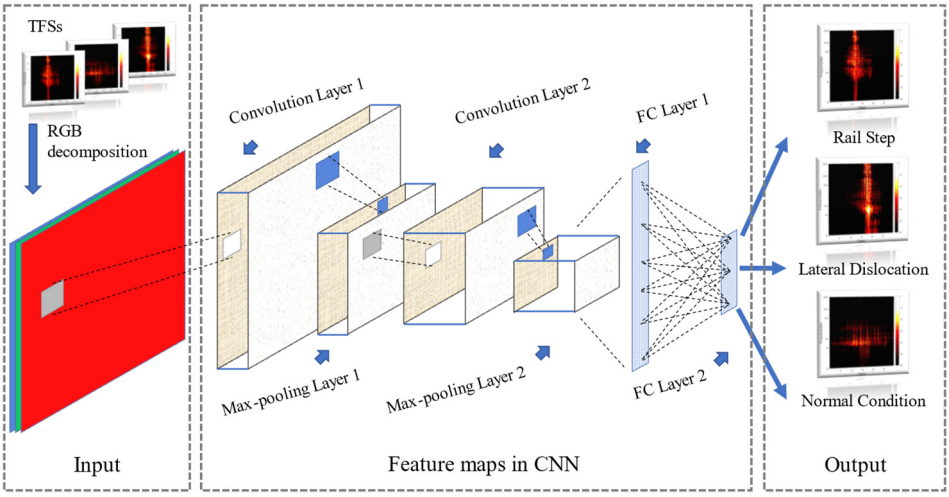


Fig. 11 Schematic diagram of the CNN-based model

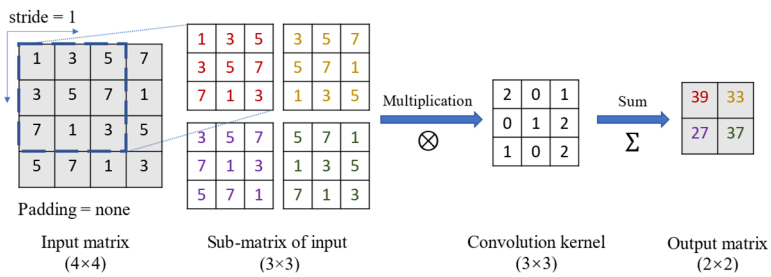
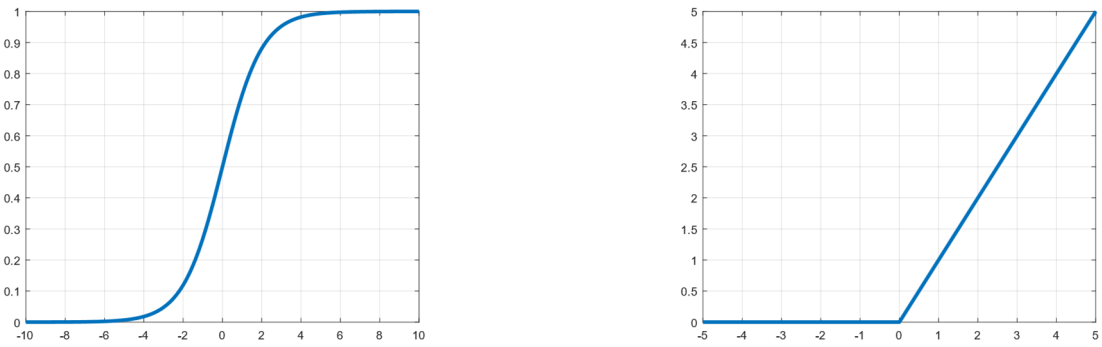


Fig. 12 One example of the convolution process



(a) Sigmoid (b) ReLU

Fig. 13 Nonlinear activation functions

nonlinear activation functions. However, the sigmoid function potentially has the "Vanishing Gradient" problem, which may cause difficulty in convergence in regard to training a deep-layer neural network. As shown in Fig. 13, unlike the sigmoid function, the gradient of ReLU is basically constant, which helps to ameliorate the convergence of a deep-layer neural network. Therefore, ReLU is adopted in the convolution layers in this study.

With the nonlinear activation function and bias, the feature map Z_k, derived from the kth convolution kernel, is defined as

$$Z_k = f(E_k + b) \tag{4}$$

where b is the bias and $f(\cdot)$ is the nonlinear activation function. It is noted that each convolution kernel will generate one feature map individually, and the output feature map combines all feature maps from different convolution kernels.

The batch normalization strategy is invoked after the introduction of nonlinearity for both convolution layers 1 and 2. Batch normalization is a straightforward and efficient method of regularization to ease the problem of internal covariate shift in the CNN model training (Ioffe and Szegedy 2015). The use of batch normalization strengthens the ability of a neural network to fit unstable distributions in every training step. As a result, not only the model generalization is raised but also the convergence speed is accelerated.

4.2.2 Max-pooling layer

To extract the most representative time-frequency features in the convolution process, a pooling layer is set after each convolution layer. There are two ways of pooling: max-pooling and average-pooling. It has been reported that max-pooling performs better than average-pooling as regards image classification (Scherer *et al*. 2010). Hence, the max-pooling operation is adopted in our CNN architecture.

In the max-pooling operation, the input feature map is divided into several non-overlapping rectangles and the maximum value within each rectangular region is exported and regarded as the most representative time-frequency features. As shown in Fig. 14, the width and height of output feature map is half of input feature map. The depth of feature map is retained, thus only one-quarter of features

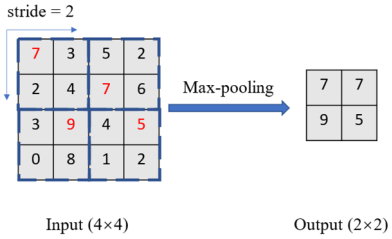


Fig. 14 The max-pooling operation

are left after the max-pooling operation, thus reduces the parameters for computation in the model training, resulting in improved computational efficiency.

4.2.3 Fully connected layer

The fully connected (FC) layers, located at the end of the CNN, target at integrating the time-frequency features to drive the classification of inputs. In this paper, the output feature map from the max-pooling layer 2 feeds into FC layer 1, where the feature maps are flattened from three-dimensional matrix to one-dimensional vector. The ReLU is attached as the nonlinear activation function in FC layer 1. And it is substituted with the Softmax function in FC layer 2 (the last FC layer) as FC layer 2 should be a classifier to turn out the probability for each class. The reason of using the Softmax function is its capacity of transferring the elements in the output vector to the probabilistic values by the normalization. Suppose that there are K classes, the output vector before the Softmax function is marked as $V = \{v_1, v_2, \ldots, v_K\}$ where each element in the vector V maps to a specific class. The Softmax function is expressed as

$$p_j = \frac{e^{v_j}}{\sum_{i=1}^{K} e^{v_j}} \tag{5}$$

where v_j is the jth element in the vector V, and $p_j \in \{p_1, p_2, \ldots, p_K\}$ is the probability of the input image belonging to the jth class. In this study, the input TFSs are categorized into three classes, that is, three states of the maglev rail joints. Hence, K is equal to 3. If p_j is the highest probability among p_1, p_2, and p_3, the jth class will be judged as the predicted state for the given input TFS. With the

output after the Softmax function, the accuracy and loss in training a batch of input TFSs are calculated by

$$\text{Accuracy} = \frac{1}{N}\sum_{m=1}^{N} l(p_m = y_m) \tag{6}$$

$$\text{Loss} = \frac{1}{N}\sum_{m=1}^{N} -[(Y_m \log(P_m) + (1 - Y_m)\log(1 - P_m)] \tag{7}$$

where N is the number of TFSs, $y_m \in Y_m$, $p_m \in P_m$. For the mth TFS, Y_m is a one-hot-coded vector, such as [0 1 0], which indicates the true probability distribution on each class. P_m indicates the predicted probability distribution, such as $[p_1\ p_2\ p_3]$. y_m reflects the exact state of a maglev rail joint, while p_m reflects the predicted state. $l(\cdot)$ is the indicator function, which equals 1 when $p_m = y_m$, otherwise equals 0.

In addition, to alleviate potential overfitting in a fully-connected deep neural network, the dropout strategy is implemented between FC layers 1 and 2. The main function of dropout is to randomly cut the connections between nodes at a specific rate. As suggested by Srivastava *et al.* (2014), the dropout rate is set to 0.5 in this study, implying that only half the nodes in FC layer 2 retain connection to FC layer 1. As a result, it blocks the co-adaptation between nodes and increases the robustness of the deep neural network.

4.3 Training of CNN-based model

Appropriate layer sizes should be determined before training the CNN model. Table 1 summarizes three types of CNN architecture which share the majority of layer sizes and exactly same improvement strategies. The three-dimensional attribute of initial input TFSs [width, height, depth] denotes [time, frequency, feature] for the three types of CNN architecture. The width and height of TFSs are both resized to 64 pixels, and the depth is the RGB values of input images. In the convolution layers, the stride is fixed to 1, the zero padding is for the input matrix, and kernels with the size of 3×3 are used. In the max-pooling layers, the stride is fixed to 2, kernels with the size of 2×2 are used. The differences of the three types of CNN architecture lie in

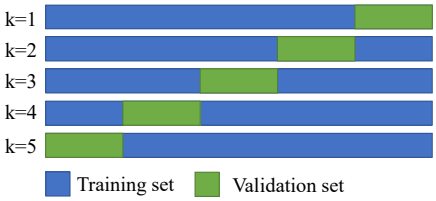
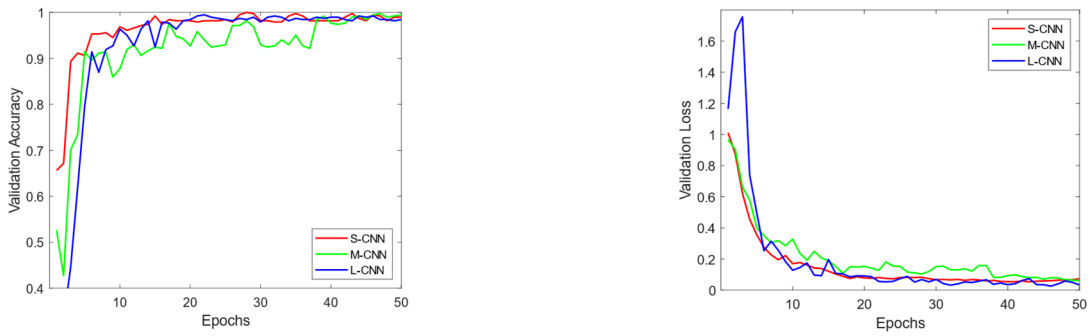


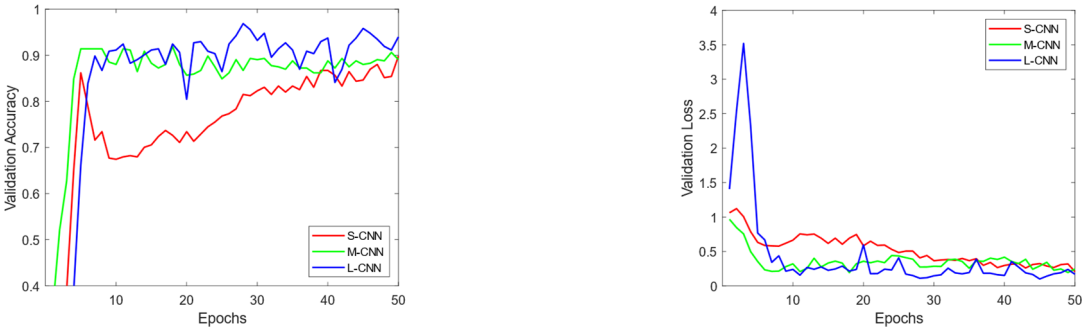
Fig. 15 Five-fold cross-validation

Table 1 Three types of CNN architecture

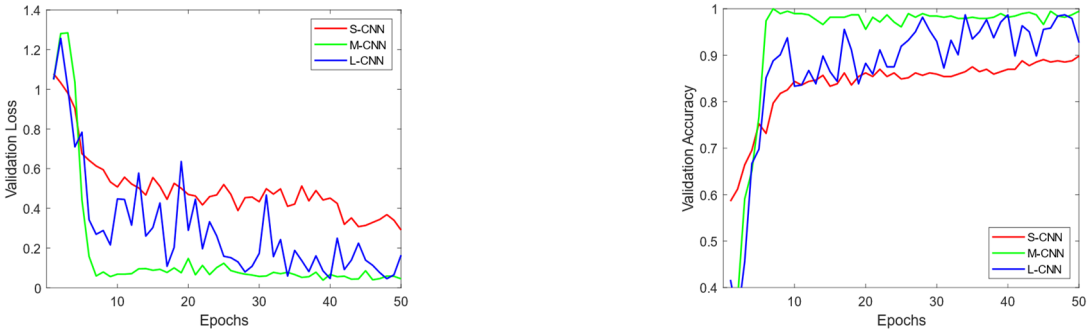
	S-CNN		M-CNN		L-CNN	
Convolution ayer 1	Initial input TFS size: [64,64,3]					
	Kernel size	4 × [3,3]	Kernel size	8 × [3,3]	Kernel size	16 × [3,3]
	Output size	[62,62,4]	Output size	[62,62,8]	Output size	[62,62,16]
	Activation function: ReLU					
Batch normalization						
Max-pooling llayer 1	Input size	[62,62,4]	Input size	[62,62,8]	Input size	[62,62,16]
	Kernel size	[2,2]	Kernel size	[2,2]	Kernel size	[2,2]
	Output size	[31,31,4]	Output size	[31,31,8]	Output size	[31,31,16]
Convolution layer 2	Input size	[31,31,4]	Input size	[31,31,8]	Input size	[31,31,16]
	Kernel size	8 × [3,3]	Kernel size	16 × [3,3]	Kernel size	32 × [3,3]
	Output size	[29,29,8]	Output size	[29,29,16]	Output size	[29,29,32]
	Activation function: ReLU					
Batch normalization						
Max-pooling layer 2	Input size	[29,29,8]	Input size	[29,29,16]	Input size	[29,29,32]
	Kernel size	[2,2]	Kernel size	[2,2]	Kernel size	[2,2]
	Output size	[14,14,8]	Output size	[14,14,16]	Output size	[31,31,32]
Fully connected layer 1	Flatten	1568	Flatten	3136	Flatten	6272
	Nodes	64	Nodes	128	Nodes	256
	Activation function: ReLU					
Dropout						
Fully connected layer 2	Nodes	3	Nodes	3	Nodes	3
	Activation function: Softmax					
Trainable parameters	101,043		403,363		1,611,843	



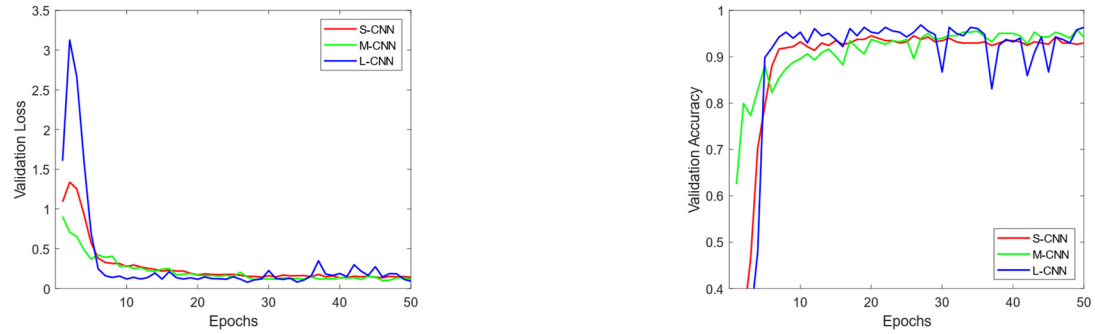
(a) k = 1



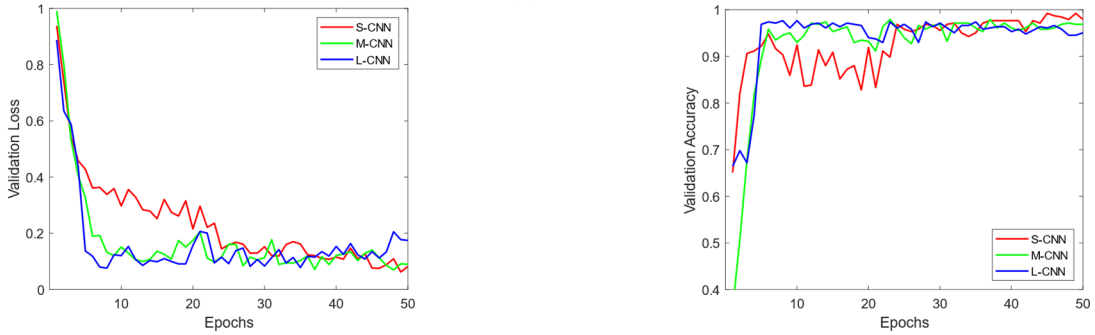
(b) k = 2



(c) k = 3



(d) k = 4



(e) k = 5

Fig. 16 Validation accuracy and loss for five runs

that the numbers of kernels are 4, 8, 16 in convolution layer 1 and 8, 16, 32 in convolution layer 2, the numbers of nodes are 64, 128, 256 in FC layer 1, and the trainable parameters are 0.1, 0.4, 1.6 (million). Thus, with the trainable parameters from low to high, the three types of CNN architecture are named as small-sized CNN (S-CNN), middle-sized CNN (M-CNN), and large-sized CNN (L-CNN), respectively. The dataset for training the three types of CNN architecture, involving the data collected in September and October 2020, is divided into the training set, validation set, and testing set. In the training process of the models, only training and validation sets are involved. The training set is used for learning the time-frequency features during the training process, while the validation set is for adjusting the model parameters to optimize the feature learning. The testing set, which is completely independent of the training and validation sets, is used for examining the effectiveness on the classification of maglev rail joint states in Section 5. The three types of CNN architecture are constructed on Keras (version 2.4.3) Application Programming Interface in the TensorFlow library written with Python 3.8. Each type of model is trained with mini-batch size of 16, optimized by Adam with the learning rate of 1×10^{-4}, for a sum of 50 epochs with 96 mini-batches per epoch. Then, the selection of the most suitable type of CNN architecture, i.e., the decision of layer sizes, is elaborated in the following.

As regards the training and validation sets, a total of 1920 TFSs are considered, which are obtained from the data collected on the rail joints J1, J2, and J3 in September 2020. To formulate the CNN model, the training and validation sets are randomized and structured to form a five-fold cross-validation (Kohavi 1995). As shown in Fig. 15, in the five-fold cross-validation, 1920 TFSs are divided into five equal parts, in which 80% of the TFSs focus on training and the remaining 20% on validation in every run of the model training. Hence, there are five parallel runs, denoted by k = 1 to k = 5, for each type of CNN architecture. To figure out an appropriate architecture of the CNN-based damage detection model, the three types of CNN architecture are evaluated using the data in the validation set in terms of accuracy and loss defined in Eqs. (6) and (7). The validation accuracy and loss of S-CNN, M-CNN, and L-CNN in five runs are shown in Fig. 16. It is observed that the convergence of validation accuracy is achieved for all three types of CNN architecture, which proves that the established CNN models enable to gradually learn the time-frequency features from TFSs with the increase of epochs. For S-CNN, the growth of accuracy and the drop of loss are both slow as shown in Figs. 16(b), (c), and (e), and even the accuracy decreases temporarily in Figs. 16(b) and (e). Moreover, the rate of convergence in S-CNN is slower than that in M-CNN and L-CNN. As for L-CNN, the accuracy and loss are both fluctuating strongly (see Figs. 16(b), (c), and (d)), and the number of trainable parameters in L-CNN is four times higher than that of M-CNN. Therefore, the damage detection model based on M-CNN is selected as the recommended model architecture.

Table 2 Results obtained by the formulated CNN-based model

Location of maglev rail joint		Predicted classification from CNN-based model			Total TFSs
		Rail step	Lateral dislocation	Normal condition	
k = 1	J1	150	10	0	160
	J2	7	147	6	160
	J3	0	10	150	160
	J4	76	1	3	80
	J5	0	0	80	80
k = 2	J1	160	0	0	160
	J2	0	145	15	160
	J3	0	8	152	160
	J4	76	4	0	80
	J5	4	6	70	80
k = 3	J1	147	13	0	160
	J2	2	146	12	160
	J3	0	14	146	160
	J4	78	1	1	80
	J5	6	1	73	80
k = 4	J1	152	3	5	160
	J2	4	145	11	160
	J3	0	13	147	160
	J4	67	1	12	80
	J5	5	3	72	80
k = 5	J1	152	1	7	160
	J2	0	159	1	160
	J3	0	12	148	160
	J4	79	1	0	80
	J5	8	5	67	80

5. Results and discussion

5.1 Model testing

In the testing stage, a total of 640 TFSs collected from five maglev rail joints (J1 to J5) in September 2020, are used to test the formulated CNN-based damage detection model. It is noted that the data collected from J4 and J5 are not involved in the model training process. The testing results will return a series of output label predicting the state of the maglev rail joints for each input TFS. Specifically, the label '0' is assigned to the TFS indicative of rail step, '1' is assigned to the TFS indicative of lateral dislocation, and '2' is assigned to the TFS indicative of normal condition. With the architecture of M-CNN, five models trained by five parallel runs are all utilized in the testing stage. Eventually, the results on the predicted classification of the five maglev rail joints are obtained as shown in Table 2.

As can be seen in Table 2, most results indicate that the rail joints J1 and J4 are under the damage condition of rail step, the rail joint J2 is under the damage condition of lateral dislocation, and the rail joints J3 and J5 are under normal condition. The prediction results are consistent with the real conditions of the five maglev rail joints. Especially, the conditions of J4 and J5 are well predicted although data collected from J4 and J5 were not used in the course of model training. The results indicate that the proposed TFS-CNN model is reliable and scalable to identify the different conditions of maglev rail joints.

5.2 Model evaluation

In order to quantitatively evaluate the model prediction performance, three indicators, including precision, recall, and F1-score, are adopted in this study. To explain these indicators, taking the damage condition of rail step as an example, all TFSs in the testing set can be predicted into two classes, marked as P (positive, rail step) and N (negative, not rail step). Compared to their actual classification, truly classified TFSs in the classes P and N are marked as TP (true positive) and TN (true negative), respectively, and falsely classified TFSs in the classes P and N are marked as FP (false positive) and FN (false negative). Then, the precision and recall are represented as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

where the precision represents the percentage of the true positive over predicted positive TFSs and the recall represents the percentage of the true positive over positive TFSs. Both the precision and recall conduct a rough inspection on the accuracy of model prediction. In addition, F1-score, designed for comprehensively considering precision and recall, is a better indicator to seek a balance between the precision and recall on classification accuracy and can be expressed as

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

The calculated precision, recall, and F1-score are shown in Fig. 17. It is found that for all three conditions, the values of precision, recall, and F1-score are greater than 83%, indicating that the formulated TFS-CNN model has an outstanding damage detection performance. Especially, the results for rail step detection are superior to the other two as the values of precision, recall, and F1-score are all higher than 92%. Although the results of lateral dislocation and normal condition are not as good as those on rail step, the values of precision, recall and F1-score are all considerably high, ranging from 83% to 95%. The prediction results of normal condition and rail step are favorably stable, while a fluctuation is observed on the prediction results of lateral dislocation. This is probably because the features shown in lateral dislocation are somewhat overlapped with the features in rail step when the amplitude of low frequency components in the lateral dislocation is small.

There is a concern as to whether the formulated damage detection model can keep satisfactory prediction performance when the condition of a maglev rail joint is changed later. To further verify the effectiveness of the TFS-CNN damage detection model, some new data, including those collected before and after replacing the loosed bolts on the rail joints J1 and J4 in October 2020, are used to train an updated model for comparison with the original model. The new training set contains TFSs converted from the data collected on the rail joints J1, J2, and J3 before 20th October 2020, and the new testing set contains TFSs from the data collected on the rail joints J1 to J5 after 20th October 2020. Both the original and updated models use the same CNN architecture, and are evaluated by using the new testing set. Table 3 provides the recall rate of each run, as well as the average recall rate, in predicting the condition of each maglev rail joint.

It is seen in Table 3 that the TFS-CNN damage detection model is competent to identify the multiple damage when damage condition is changed, as the majority of recall rates is over 80% regardless of using the original or updated model. The average recall rates for the rail joints

Table 3 The recall rates of maglev rail joints after replacing looseness bolts in J1 and J4

	Location	k = 1	k = 2	k = 3	k = 4	k = 5	Averaged
With original model	J1	96.4%	92.9%	88.4%	85.7%	93.8%	91.4%
	J2	84.8%	89.3%	89.3%	98.2%	96.4%	91.6%
	J3	96.4%	95.5%	89.3%	88.4%	92.0%	92.3%
	J4	91.3%	96.3%	80.0%	61.3%	86.3%	83.0%
	J5	96.9%	96.9%	79.7%	93.8%	82.8%	90.0%
	Location	k = 1	k = 2	k = 3	k = 4	k = 5	Averaged
With updated model	J1	98.2%	94.6%	86.6%	100.0%	86.6%	93.2%
	J2	94.6%	94.6%	99.1%	84.8%	90.2%	92.7%
	J3	93.8%	95.5%	96.4%	88.4%	91.1%	93.0%
	J4	97.5%	86.3%	83.8%	90.0%	92.5%	90.0%
	J5	100.0%	92.2%	95.3%	95.3%	90.6%	94.7%

J1 and J4, where the condition is changed from rail step to normal condition, are 91.4% and 83.0%, respectively. It verifies that the TFS-CNN model can adapt to the change in joint condition. The accuracy of prediction is improved after model updating because of the expansion of training data, where all average recall rates obtained using the updated model are more than 90% which is higher than that obtained by the original model. Especially, the average recall rates for the rail joints J1 and J4 are increased by 1.8% and 7.0%, respectively. This indicates that the TFS-CNN model updating can beneficially ameliorate prediction accuracy even when damage condition is changed.

5.3 Comparison with other neural networks

To confirm the superiority of the TFS-CNN model for multiple damage detection of maglev rail joints, two other neural network models are also formulated for comparison: a CNN model trained using power spectrum densities (PSDs) as its input and a traditional neural network (NN) model using TFSs as its input. In the training process, both models use the data collected in September 2020 and the five-fold cross validation is adopted. In the testing stage, only data collected on the rail joints J1, J2 and J3 are considered. The damage detection capabilities of these two neural networks are evaluated with the three indicators, and the results are compared with those obtained by the CNN model with TFS input (TFS-CNN).

5.3.1 CNN model with PSD input

The performance indicators obtained from the CNN model with TFS input and the CNN model with PSD input are shown in Fig. 18. It is illustrated that the performances of the two models for identifying the state of lateral dislocation are almost consistent as the obtained values of precision, recall, and F1-score range from 85% to 95%. However, the performance of the CNN model with TFS input for identifying rail step and normal condition of the rail joints is better than that of the CNN model with PSD input. For example, the values of the three indicators (93.7%, 93.6%, and 93.6% for precision, recall, and F1-score) for normal condition obtained by the CNN model

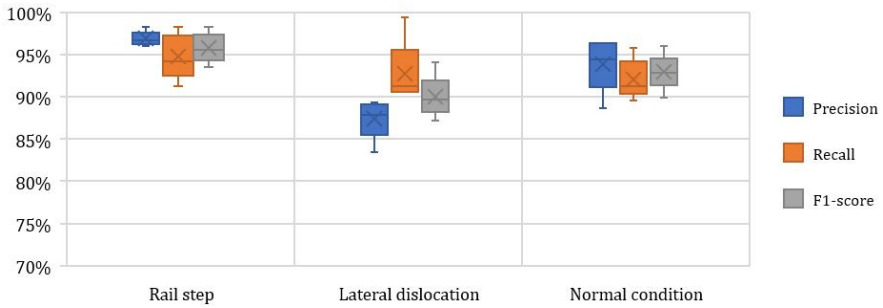


Fig. 17 The precision, recall, and F1-score for detection of different conditions

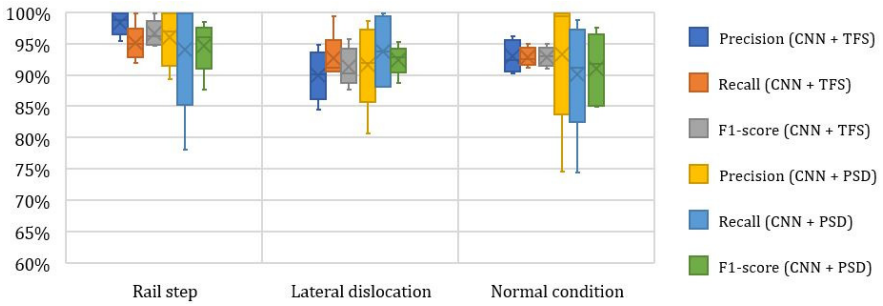


Fig. 18 The precision, recall, and F1-score of CNN models trained respectively with TFS and PSD input

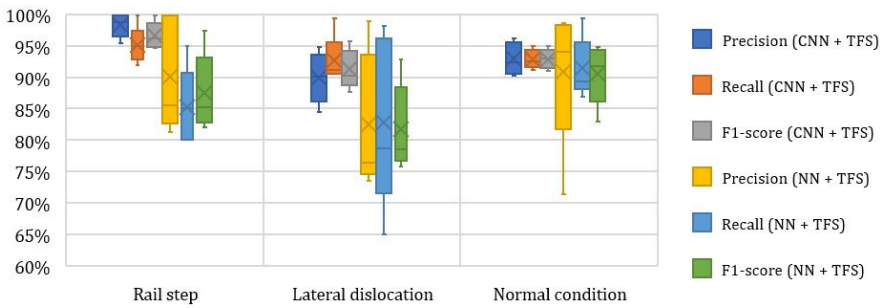


Fig. 19 The precision, recall, and F1-score of CNN and NN models both trained with TFS input

with TFS input are larger than those obtained by the CNN model with PSD input (93.6%, 92.6%, and 92.7% for precision, recall, and F1-score). In addition, the results from the CNN model with PSD input are less stable in that the standard deviations of the precision, recall, and F1-score (0.098, 0.084, and 0.051) are larger than those (0.023, 0.014, and 0.014) from the CNN model with TFS input. Similar observation is made on the values of the three indicators for rail step. In summary, the performance of the CNN model with TFS input (TFS-CNN) is better than the CNN model with PSD input.

5.3.2 Traditional neural network with TFS input

In the traditional neural network (NN), the TFS, which is the same as used in the TFS-CNN model, is initially flattened to a series of one-dimensional vector as input to the neural network. The NN model consists of two hidden layers with 128 and 64 neurons and ends up with a 3-neurons layer. Activation functions are also adopted at each layer, with Softmax for the last layer and ReLU for the other layers. The NN model is trained with the batch size of 16 and the learning rate of 1×10^{-4}.

The values of precision, recall, and F1-score obtained by the CNN model with TFS input (TFS-CNN) and the NN model with TFS input are shown in Fig. 19. As can be seen, the minimum and average values of precision, recall, and F1-score obtained from the NN model are much lower than those from the CNN model. In addition, the standard deviations of the three indicators for the NN model are much higher than those from the CNN model. Overall, the traditional neural network model enables the prediction of damage detection, but its robustness is far worse than the TFS-CNN model. This is probably due to the lack of convolutional and max-pooling operations in the traditional neural network for extraction of deeper and hidden features in TFSs.

To facilitate comparison, the average values and standard deviations of the three indicators obtained from the CNN model with TFS input, the CNN model with PSD input, and the NN with TFS input are listed in Table 4. Among them, generally, the highest average value and the lowest standard deviation are seen in the CNN model with TFS input, indicating that the TFS-CNN model has advantages of higher accuracy and scalability for multiple damage detection of maglev rail joints.

6. Conclusions

In this study, a TFS-CNN model has been developed to realize multiple damage detection of maglev rail joints. By means of STFT, the collected raw accelerations from an online monitoring system are transformed into TFSs which contain the time-frequency spectral features of maglev rail joints. A CNN-based model is then constructed by using the dataset of TFSs, which is also updatable with the cumulation of collected data and the awareness of real condition of maglev rail joints. The CNN configuration is appropriately designed through comparing the performance of different CNN architectures when being fed with validation data.

The capability of the proposed TFS-CNN damage detection model was verified by using the rail acceleration response data collected during the trial runs of a maglev train on a test line where the devised online monitoring system was deployed. In the field tests, five maglev rail joints with three different conditions: rail step caused by bolt looseness, lateral dislocation caused by installation error, and normal condition, were identified by the proposed method. The field test results confirm that: (i) by use of rail acceleration measurements at two ends of the two F-type rail, the damage condition of maglev rail joints can be reliably identified by the proposed TFS-CNN model; (ii) when directly using the frequency spectrum to train the CNN model, the precision, recall, and F1-score are reduced and the identification results are less stable; and (iii) in comparison with the traditional neural network trained with the same TFS input, the TFS-CNN model is more robust.

Acknowledgments

The research described in this paper was supported by a grant (RIF) from the Research Grants Council of the Hong Kong Special Administrative Region (SAR), China (Grant No. R-5020-18), a grant from the National Natural Science Foundation of China (Grant No. U1934209) and Wuyi University's Hong Kong and Macao Joint Research and Development Fund (Grants No. 2019WGALH15 and 2019WGALH17). The authors would also like to appreciate the funding support by the Innovation and Technology Commission of the Hong Kong SAR Government to the

Table 4 The indicators of three neural networks

Indicators		CNN with TFS input			CNN with PSD input			NN with TFS input		
		Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Rail step	Avg.	0.983	0.951	0.966	0.960	0.941	0.947	0.901	0.853	0.875
	Std.	0.017	0.026	0.019	0.040	0.085	0.037	0.081	0.055	0.055
Lateral dislocation	Avg.	0.899	0.927	0.913	0.916	0.937	0.924	0.825	0.829	0.817
	Std.	0.036	0.034	0.028	0.061	0.050	0.021	0.097	0.120	0.062
Normal condition	Avg.	0.929	0.929	0.929	0.933	0.901	0.910	0.908	0.914	0.905
	Std.	0.023	0.014	0.014	0.098	0.084	0.051	0.101	0.043	0.042
Mean Avg.		0.937	0.936	0.936	0.936	0.926	0.927	0.878	0.865	0.866
Mean Std.		0.025	0.025	0.020	0.066	0.073	0.037	0.093	0.073	0.053

Hong Kong Branch of Chinese National Rail Transit Electrification and Automation Engineering Technology Research Center (Grant No. K-BBY1) and by the National Natural Science Foundation of China to the Maglev Transportation Engineering R&D Center (Grant No. 52072269).

References

Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M. and Inman, D.J. (2017), "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks", *J. Sound Vib.*, **388**, 154-170. https://doi.org/10.1016/j.jsv.2016.10.043

Alberts, T.E., Hanasoge, A.M. and Omran, A.M. (2011), "On the influence of structural flexibility on feedback control stability for magnetically suspended vehicles", *J. Dyn. Syst. Measur. Control*, **133**(5), 051010. https://doi.org/10.1115/1.4003802

Cha, Y.J., Choi, W., Suh, G., Mahmoudkhani, S. and Büyüköztürk, O. (2017), "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types", *Comput.-Aided Civil Infrastr. Eng.*, **33**(9), 731-747. https://doi.org/ 10.1111/mice.12334

Chen, J.W., Liu, Z.G., Wang, H.R., Núñez, A. and Han, Z.W. (2018), "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network", *IEEE Transact. Instrument. Measur.*, **67**(2), 257-269. https://doi.org/10.1109/TIM.2017.2775345

Dangre, H. (2019), "A review on insulated rail joints (IRJ) failure analysis", *Int. J. Adv. Res. Publicat.*, **3**(1), 5-9.

Deng, L., Chu, H.H., Shi, P., Wang, W. and Kong, X. (2020), "Region-based CNN method with deformable modules for visually classifying concrete cracks", *Appl. Sci.*, **10**(7), 2528. https://doi.org/10.3390/app10072528

Duan, Y.F., Chen, Q.Y., Zhang, H.M., Yun, C.B., Wu, S.K. and Zhu, Q. (2019), "CNN-based damage identification method of tied-arch bridge using spatial-spectral information", *Smart Struct. Syst., Int. J.*, **23**(5), 507-520. https://doi.org/10.12989/sss.2019.23.5.507

Dung, C.V., Sekiya, H., Hirano, S., Okatani, T. and Miki, C. (2019), "A vision-based method for crack detection in gusset plate welded joints of steel bridges using deep convolutional neural networks", *Automat. Constr.*, **102**, 217-229. https://doi.org/10.1016/j.autcon.2019.02.013

Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R. and De Schutter, B. (2016), "Deep convolutional neural networks for detection of rail surface defects", *Proceedings of 2016 International Joint Conference on Neural Networks* (*IJCNN*), Vancouver, Canada, July.

Feng, J., Li, F.M., Lu, S.X., Liu, J.H. and Ma, D.Z. (2017), "Injurious or noninjurious defect identification from MFL images in pipeline inspection using convolutional neural network", *IEEE Transact. Instrument. Measur.*, **66**(7), 1883-1892. https://doi.org/10.1109/TIM.2017.2673024

Fujie, J. (1989), "Current status of EDS system in Japan", *Proceedings of 11th International Conference on Magnetically Levitated Systems and Linear Drives*, Yokohama, Japan, July.

Gibert, X., Patel, V.M. and Chellappa, R. (2017), "Deep multitask learning for railway track inspection", *IEEE Transact. Intell. Transport. Syst.*, **18**(1), 153-164. https://doi.org/10.1109/TITS.2016.2568758

Goodall, R.M. (2008), "Generalised design models for EMS maglev", *Proceedings of 20th International Conference on Magnetically Levitated Systems and Linear Drives*, San Diego, CA, USA, December.

Huang, J.Y., Wu, Z.W., Shi, J., Gao, Y. and Wang, D.Z. (2018), "Influence of track irregularities in high-speed Maglev transportation systems", *Smart Struct. Syst., Int. J.*, **21**(5), 571-582. https://doi.org/10.12989/sss.2018.21.5.571

Ioffe, S. and Szegedy, C. (2015), "Batch normalization: accelerating deep network training by reducing internal covariate shift", *Proceedings of 32nd International Conference on Machine Learning*, Lile, France, July.

Kabo, E., Nielsen, J.C.O. and Ekberg, A. (2006), "Prediction of dynamic train–track interaction and subsequent material deterioration in the presence of insulated rail joints", *Vehicle Syst. Dyn.*, **44**, 718-729. https://doi.org/10.1080/00423110600885715

Kim, K.J, Han, J.B., Han, H.S. and Yang, S.J. (2015), "Coupled vibration analysis of Maglev vehicle-guideway while standing still or moving at low speeds", *Vehicle Syst. Dyn.*, **53**(4) 587-601. http://dx.doi.org/10.1080/00423114.2015.1013039

Kim, J.J., Kim, A.R. and Lee, S.W. (2020), "Artificial neural network-based automated crack detection and analysis for the inspection of concrete structures", *Appl. Sci.*, **10**(22), 8105. https://doi.org/10.3390/app10228105

Kohavi, R. (1995), "A study of cross-validation and bootstrap for accuracy estimation and model selection", *Proceedings of 14th International Joint Conference on Artificial Intelligence* (*IJCAI '95*), San Francisco, CA, USA, August.

Li, X.Y., Liu, Y.Z., Zhou, W.W. and Wu, J. (2015), "Research on the influence on running vehicle from non-coplanar disturbance of four magnetic pole surfaces of middle-low speed maglev train tracks", *Proceedings of 2015 Seventh International Conference on Measuring Technology and Mechatronics Automation*, Nanchang, China, June.

Li, Y.J., Yu, P.C., Zhou, D.F. and Li, J. (2018), "Magnetic flux feedback strategy to suppress the gap fluctuation of low-speed maglev train caused by track steps", *Proceedings of 37th Chinese Control Conference* (*CCC*), Wuhan, China, July.

Li, H.T., Xu, H.Y., Tian, X.D., Wang, Y., Cai, H.Y., Cui, K.R. and Chen, X.D. (2020a), "Bridge crack detection based on SSENets", *Appl. Sci.*, **10**(12), 4230. https://doi.org/10.3390/app10124230

Li, X.Q., Zhai, M.D., Li, X.L. and Dong, W.H. (2020b), "Research on suppression strategy for track dislocation interference in medium-low speed maglev train", *Proceedings of 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference* (*ITNEC*), Chongqing, China, June.

Lin, Y.Z., Nie, Z.H. and Ma, H.W. (2017), "Structural damage detection with automatic feature-extraction through deep learning", *Comput.-Aided Civil Infrastr. Eng.*, **32**(12), 1025-1046. https://doi.org/10.1111/mice.12313

Luo, L.X., Feng, M.Q., Wu, J.P. and Leung, R.Y. (2019), "Autonomous pothole detection using deep region-based convolutional neural network with cloud computing", *Smart Struct. Syst., Int. J.*, **24**(6), 745-757. https://doi.org/10.12989/sss.2019.24.6.745

Ma, Z.R., Gao, L., Zhong, Y.L., Ma, S. and An, B.L. (2020), "Arching detection method of slab track in high-speed railway based on track geometry data", *Appl. Sci.*, **10**(19), 6799. https://doi.org/10.3390/app10196799

Masada, E. (1993), "Development of maglev transportation in Japan: present state and future prospects", *Proceedings of 13th International Conference on Magnetically Levitated Systems and Linear Drives*, Argonne, IL, USA, May.

Marino, F., Distante, A., Mazzeo, P.L. and Stella, E. (2007), "A real-time visual inspection system for railway maintenance: automatic hexagonal-headed bolts detection", *IEEE Transact. Syst. Man Cybernet. Part C* (*Applications and Reviews*), **37**(3), 418-428. https://doi.org/10.1109/TSMCC.2007.893278

Ohtsuka, T. and Iguchi, M. (1982), "Maglev dynamics and ride quality: past, present and future", *Proceedings of 2nd International Seminar on the Superconductive Magnetic Levitated Train*, Miyazaki, Japan, November.

Oregui, M., Li, S., Núñez, A., Li, Z., Carroll, R. and Dollevoet, R. (2017), "Monitoring bolt tightness of rail joints using axle box acceleration measurements", *Struct. Control Health Monitor.*, **24**(2), e1848. https://doi.org/10.1002/stc.1848

Park, J.H., Kim, T.H. and Kim, J.T. (2015), "Image-based bolt-loosening detection technique of bolt Joint in steel bridges", *Proceedings of 6th International Conference on Advances in Experimental Structural Engineering*, Urbana-Champaign, IL, USA.

Pham, H.C., Ta, Q.B., Kim, J.T., Ho, D.D., Tran, X.L. and Huynh, T.C. (2020a), "Bolt-loosening monitoring framework using an image-based deep learning and graphical model", *Sensors*, **20**(12), 3382. https://doi.org/10.3390/s20123382

Pham, M.T., Kim, J.M. and Kim, C.H. (2020b), "Accurate bearing fault diagnosis under variable shaft speed using convolutional neural networks and vibration spectrogram", *Appl. Sci.*, **10**(18), 6385. https://doi.org/10.3390/app10186385

Sato, Y., Matsuura, A., Miura, S. and Satoh, Y. (1985), "Development of guideway for maglev", *Proceedings of 7th International Conference on Maglev Transport'85*, Tokyo, Japan, September.

Scherer, D., Müller, A. and Behnke, S. (2010), "Evaluation of pooling operations in convolutional architectures for object recognition", *Proceedings of 20th International Conference on Artificial Neural Networks* (*ICANN*), Thessaloniki, Greece, September.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014), "Dropout: a simple way to prevent neural networks from overfitting", *J. Mach. Learn. Res.*, **15**(1), 1929-1958.

Sung, H.K., Jho, J.M., Bae, D.K., Rho, K.S., Lee, J.M., Yoo, M.H. and Nam, Y.Y. (2006), "A fuzzy based treatment to reduce air-gap disturbance at the rail joints with step-wise rail joint", *Proceedings of 19th International Conference on Magnetically Levitated Systems and Linear Drives*, Dresden, Germany, September.

Teng, Z.Q., Teng, S., Zhang, J.Q., Chen, G.F. and Cui, F.S. (2020), "Structural damage detection based on real-time vibration signal and convolutional neural network", *Appl. Sci.*, **10**(14), 4720. https://doi.org/10.3390/app10144720

Verstraete, D., Ferrada, A., Droguett, E.L., Meruane, V. and Modarres, M. (2017), "Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings", *Shock Vib.*, 5067651. https://doi.org/10.1155/2017/5067651

Wu, D., Sun, X., Chang, Y., Xu, W., Huang, J., Wu, Z. and Wang, D. (2019), "The temperature effect analysis of high-speed maglev transit", *Proceedings of 4th International Conference on Automatic Control and Mechatronic Engineering* (*ACME 2019*), Chongqing, China, May.

Yau, J.D. (2009), "Response of a maglev vehicle moving on a series of guideways with differential settlement", *J. Sound Vib.*, **324**, 3-5. https://doi.org/10.1016.jsv.2009.02.031

Ye, X.W., Jin, T. and Yun, C.B. (2019), "A review on deep learning-based structural health monitoring of civil infrastructures", *Smart Struct. Syst., Int. J.*, **24**(5), 567-585. https://doi.org/10.12989/sss.2019.24.5.567

Zhang, L. and Huang, J.Y. (2018), "Thermal effect on dynamic performance of high-speed maglev train/guideway system", *Struct. Eng. Mech., Int. J.*, **68**(4), 459-473. https://doi.org/10.12989/sem.2018.68.4.459

Zhang, D.K., Gao, S.B., Yu, L. and Zhan, D. (2017), "Dynamic detection method of medium-low speed maglev F-track seams based on machine vision", *CES Transact. Electr. Mach. Syst.*, **1**(4), 343-353. https://doi.org/10.23919/TEMS.2017.8241355

Zhang, Y., Sun, X.W., Loh, K.J., Su, W.S., Xue, Z.G. and Zhao, X.F. (2020), "Autonomous bolt loosening detection using deep learning", *Struct. Health Monitor.*, **19**(1), 105-122. https://doi.org/10.1177/1475921719837509

Zhao, X.F., Zhang, Y. and Wang, N.N. (2019), "Bolt loosening angle detection technology using deep learning", *Struct. Control Health Monitor.*, **26**(1), e2292. https://doi.org/10.1002/stc.2292

Zheng, R.H., Xiong, C., Deng, X.B., Li, Q.S. and Li, Y. (2020), "Assessment of earthquake destructive power to structures based on machine learning methods", *Appl. Sci.*, **10**(18), 6210. https://doi.org/10.3390/app10186210

Zhong, K.F., Teng, S., Liu, G., Chen, G.F. and Cui, F.S. (2020), "Structural damage features extracted by convolutional neural networks from mode shapes", *Appl. Sci.*, **10**(12), 4247. https://doi.org/10.3390/app10124247

Zhou, D.F., Yu, P.C., Wang, L.C. and Li, J. (2017), "An adaptive vibration control method to suppress the vibration of the maglev train caused by track irregularities", *J. Sound Vib.*, **408**, 331-350. https://doi.org/10.1016/j.jsv.2017.07.037

HJ