

# Structural damage detection in presence of temperature variability using 2D CNN integrated with EMD

Smriti Sharma<sup>a</sup> and Subhamoy Sen\*

*i4S Laboratory, Indian Institute of Technology Mandi, Mandi, HP, India*

*(Received July 22, 2021, Revised December 1, 2021, Accepted December 17, 2021)*

**Abstract.** Traditional approaches for structural health monitoring (SHM) seldom take ambient uncertainty (temperature, humidity, ambient vibration) into consideration, while their impacts on structural responses are substantial, leading to a possibility of raising false alarms. A few predictors model-based approaches deal with these uncertainties through complex numerical models running online, rendering the SHM approach to be compute-intensive, slow, and sometimes not practical. Also, with model-based approaches, the imperative need for a precise understanding of the structure often poses a problem for not so well understood complex systems. The present study employs a data-based approach coupled with Empirical mode decomposition (EMD) to correlate recorded response time histories under varying temperature conditions to corresponding damage scenarios. EMD decomposes the response signal into a finite set of intrinsic mode functions (IMFs). A two-dimensional Convolutional Neural Network (2DCNN) is further trained to associate these IMFs to the respective damage cases. The use of IMFs in place of raw signals helps to reduce the impact of sensor noise while preserving the essential spatio-temporal information less-sensitive to thermal effects and thereby stands as a better damage-sensitive feature than the raw signal itself. The proposed algorithm is numerically tested on a single span bridge under varying temperature conditions for different damage severities. The dynamic strain is recorded as the response since they are frame-invariant and cheaper to install. The proposed algorithm has been observed to be damage sensitive as well as sufficiently robust against measurement noise.

**Keywords:** Convolutional Neural Network (CNN); damage detection; Deep Learning (DL); Empirical Mode Decomposition (EMD); Structural Health Monitoring (SHM)

## 1. Introduction

The primary objective of Structural Health Monitoring (SHM) is to determine the occurrence, location, and severity of damage (Sohn *et al.* 1996) promptly and precisely. In this regard, vibration based SHM has proved its merits as one of the reliable and efficient approach to ensure the safety of structures in operation. With such approaches, the damage is mostly demarcated as deterioration in material, geometrical, or boundary properties having adverse effects on the structural stiffness and dynamic properties. Eventually, the damage is expected to leave its signature on the vibrational response. Nevertheless, not only the damage but varying

---

\*Corresponding author, Ph.D., E-mail: subhamoy@iitmandi.ac.in

<sup>a</sup> Ph.D. Student, E-mail: d17007@students.iitmandi.ac.in

environmental conditions, especially temperature, are observed to have a substantial impact on the vibrational response (Cornwell *et al.* 1999, Alampalli 2000). Isolating the damage-specific changes in response from the overall change is, however, a major challenge for real-life SHM (Kullaa 2009).

However, contrastingly insignificant numbers of research considered incorporation of these environmental impacts within the SHM approach. As a consequence, false alarms (positive or negative) may be signaled since the environmental effects, if not accounted for, may mask the damage-induced changes in the response. Without adopting sufficient measures to achieve robustness against environmental uncertainty, no SHM techniques can be considered reliable. This can be one of the major reasons that the application of the existing SHM techniques is contrastingly poor compared to the staggering progress in their development.

Among the several ambient uncertainties affecting SHM outcome, the present study concentrates on temperature effects only. Ambient variability in temperature has been identified as one of the major aspects that influences SHM decisions (Cornwell *et al.* 1999, Alampalli 2000) identifies that temperature-induced variations in vibrational properties may even surpass alterations caused by damage of medium severities. To counter this, most model based SHM approaches take a basis on a detailed numerical model that includes the temperature dependence of the material properties (majorly elasticity) (Glisic *et al.* 2011, Kromanis *et al.* 2016). The impact of thermal stresses on geometric stiffness also plays a major role in altering dynamic properties, discussed in very few studies (Alampalli 2000, Sharma and Sen 2021a). In reality, this geometric component of structural stiffness is very much case-specific and majorly depend on the structural configuration, boundary properties, temperature profile, etc. Further, thermal effects may cause supports to move altering the structural configuration itself (Kromanis *et al.* 2016). The inclusion of all such aspects in the predictor model may at times be complex and compute intensive.

Traditionally, damage in a structure is identified by monitoring typical damage-sensitive features (e.g., acceleration, strain time histories, modal frequency or damping, etc.) with/without a supporting predictor model replicating its dynamics. However, for most cases, these features are not robust to the operational variability and/or varying temperature conditions. This motivated researchers to opt for either more “intelligent features”, or more “intelligent approaches” that are robust and can identify damage-induced changes undeterred by the ambient variability.

With the introduction of cheap sensor technology and computational resources, Machine learning (ML)-based algorithms have come up as an efficient alternative to this cause (Sun *et al.* 2020). With ML, several smart and robust features have been proposed to detect an anomaly in the structure in presence of environmental variabilities. ML-based SHM algorithms are also observed to manifest robustness against operational variability through the employment of an enriched training database. The existing approaches have majorly remained either supervised or unsupervised depending on the problem. While with an unsupervised algorithm, the damage occurrence can be detected at much ease by sensing alteration in a damage-sensitive feature in the response, localization and/or quantification may need supervised training using rich archives of response pertaining to all possible damage cases. Such rigorous training, in turn, expects access to responses pertaining to all possible damage cases specific to the structure, which is neither achievable nor pragmatic. While access to undamaged response data is not a problem, the damaged responses are typically simulated under different damage scenarios using a predictor model, and collectively they are employed to train a network to correlate structural response to corresponding damage cases (Zhou *et al.* 2011, Yarnold and Moon 2015, Jin *et al.* 2016, Weinstein *et al.* 2018).

Overall, ML-based SHM approaches attempt two main steps in consequence: extraction of a damage-sensitive feature and subsequently classifying them to corresponding damage cases. In this regard, ML-based SHM approaches traditionally relied on “hand-crafted” features (e.g., frequencies, mode shapes, frequency response functions (FRF), modal damping, etc.) manually chosen and extracted from the raw signal well in advance. Further, a classifier is employed to map these features to corresponding damage cases. Numerous combinations of features and classifiers have been studied in the literature in search of an optimal choice for ML-based SHM. As it happens, the success of such an approach depends on the sensitivity of the chosen feature as well as on the efficacy of the classifier. As a consequence, most often, it is experienced that an algorithm that seemed efficient for a particular structure is not necessarily a good option for the others. The variation in performance can be attributed to the manual selection process of the features, failing to ensure a proper selection always.

Recent studies have demonstrated the relatively superior performance of Deep Learning (DL)-based techniques compared to the traditional ML-based SHM approaches. Unlike parametric methods, these approaches do not look for features of physical significance. Instead, the extracted features are mostly numerical and therefore abstract, yet sensitive to the damage classes. This alleviates the requirement to understand the structural behavior and select a robust feature through which damage can be classified. CNN is one such approach that automates the feature selection purging the manual feature selection process.

Typically, CNN merges feature extraction and classification together in a single block, getting rid of human selection bias. CNN, in its one/two-dimensional version, has found much application in the recent SHM research and proved to be efficient and robust. Yet, CNN has never been tested for its capability to achieve robustness against environmental uncertainty. Developed with an aim to deal with image processing problems, CNN has been modified to handle 1D time-series array (Abdeljaber *et al.* 2018, Sharma and Sen 2020). Presenting multiple time series arrays recorded from multiple sensors under the guise of images, major potentials of CNN can be exploited. This article, therefore, attempts a 2D-CNN dealing with 2D response data to detect structural damage in the presence of temperature variations.

Of course, the efficiency of the CNN-based SHM approach depends on the type and quality of the data being handled. The selection of a good data source can in fact, improve the detection ability. It is challenging to interpret structural health from the noisy ambient vibration recorded during operation. Yet, for real-life structures, the ambient response is most often the only measurable response. Such responses, measured under an uncontrolled environment, are often noise-contaminated, non-stationary, and/or non-linear. Direct employment of such data may lead to a sub-optimal solution with a CNN network conditioned on stationary and linear signals and therefore calls for a pre-processing.

Further, recent studies observe strain as a better damage-sensitive response compared to acceleration (dos Santos *et al.* 2015, Xia *et al.* 2017). This motivates this study to replace typical acceleration with strain as the measured response for the CNN to classify. For pre-processing of the input data, Empirical Modal Decomposition (EMD) technique is proposed in this study that decomposes the raw signal into finite elementary orthogonal components, termed as Intrinsic Mode Functions (IMF), that contains all essential characteristics sensitive to damage while being free from noises to some extent.

In the following, state of the art with ML-based SHM research has been discussed in Section 2 followed by a brief description of the methods involved: i.e., EMD (Section 3.1) and CNN (Section 3.2) approaches in process of detailing the proposed approach in Section 3. The method

has been numerically validated in Section 4 on a single span numerical bridge model that includes temperature-induced effects like thermal expansion and bowing encompassing their impacts on geometric stiffness. The detailed discussion on thermal effects on structural stiffness is, however, excluded in the article and the readers are requested to follow (Sharma and Sen 2021a, [b](#)) where the same has been discussed elaborately.

## 2. State of the art

The generic approach for most ML-based SHM approaches is to follow two consequent steps:

- i) deciding on a feature that is robust yet sensitive to damage and subsequent extraction of the same from the measured signal, followed by,
- ii) classification of these features among a set of predefined damage scenarios.

The literature in this aspect contains several parametric as well as non-parametric approaches that employ different strategies to use the information embedded in the measured data. The non-parametric approaches attempt for a damage-sensitive feature that can indicate a presence of damage in the system. Examples of such methods include simple statistical methods, nonlinear auto-regressive with exogenous inputs (NARX) neural networks, principal component analysis, wavelet transform, auto-regressive modelling, self-organizing maps Tibaduiza *et al.* 2013 etc. For the classification, artificial neural networks (ANN), fuzzy neural networks, probabilistic neural network (PNNs), and support vector machines (de Oliveira *et al.* 2018, Sharma and Sen 2020) have proved their efficacy for complicated SHM problems.

Eventually, the selection of the features and the classifiers dictate the accuracy and practicality of the employed ML-based SHM algorithm, and therefore should ideally be chosen with care. Handcrafted features may induce human bias in the employed SHM algorithm, which can sometimes be detrimental to overall accuracy. Further, not always a single feature can be sensitive for all possible damage cases. Yet, no thumb rule exists to help select the features or the classifiers, posing the major problem for ML-based SHM and rendering the approach to be case-specific. Further, the disjointed feature extraction and classification steps collectively make the ML-based SHM approaches compute-intensive and at times, not economical or slow. DL-based approaches like CNN have emerged as a possible breakthrough in this attempt. Rigorous comparison between classical neural network and deep learning techniques for SHM has been established DL to be a consistently superior approach in several articles (Sharma and Sen 2020, 2021b).

Inspired by the structure and operational approach of the human visual cortex, CNN was introduced by LeCun *et al.* (1989). It is an efficient approach that combines extraction of abstract feature and subsequent classification in a single learning block (Avci *et al.* 2017, Abdeljaber *et al.* 2018). Within CNN, two layers: convolution and pooling, are alternatively stacked in order to perform the following processes recursively:

- (1) The convolutional layer extracts a set of translation invariant local features (denoted as feature map) through convolution of a set of local filters with the raw signal,
- (2) The following pooling layer employs a sliding window on the convoluted signal to extract the features of higher strength. This also down-samples the feature maps through a patch wise summarizing.

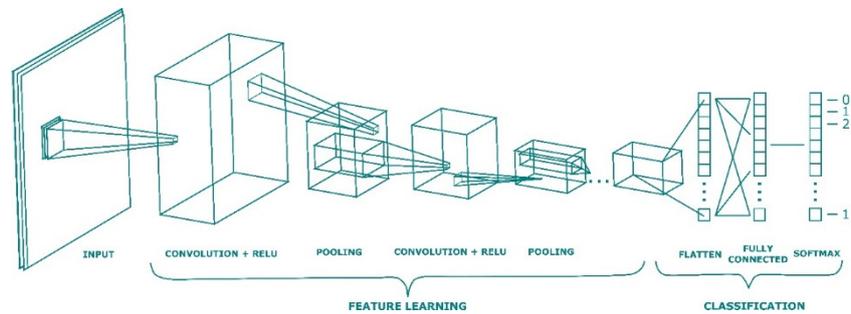


Fig. 1 2D convolutional neural network

This recursive feature extraction and data compression process is repeated over several stacks of convolutional and pooling layers prior to a classification layer that classifies the compressed feature map among different damage scenarios/classes. Cross entropy and softmax are typical choices for loss and classifier functions, respectively. It should be noted that CNN does not depend on user-supplied features, rather extracts its own abstract features. A schematic representation of a 2D CNN with multi-layer image input has been presented in Fig. 1.

CNN was originally introduced to handle image or video processing and pattern recognition in two or three-dimensional vision data. Because of the unique network architecture and feature-free classification approach, CNN has seen applications in various vision-based problems (Krizhevsky *et al.* 2012, LeCun *et al.* 1989). This approach can isolate unique features deeply embedded even within complex and uncorrelated signals, which otherwise might remain imperceptible with shallow networks or expert human investigators. Although not developed targeting classification of time series and sequential data, the astounding similarity in problem types motivated the researchers to employ CNN for problems like speech recognition (Abdel-Hamid *et al.* 2012), fault detection in engine (Ince *et al.* 2016), natural language processing (Duque *et al.* 2019), electrocardiogram classification (ECG) beats (Kiranyaz *et al.* 2016), and finally structural damage detection (Abdeljaber *et al.* 2018) etc. Such studies employed a 1D version of CNN achieving higher detection accuracy even with minimal training data (Abdeljaber *et al.* 2017, Avci *et al.* 2017).

The application of CNN in SHM research is, however, a relatively less-explored topic. The pertinent literature has been reviewed in the works of Zhao *et al.* (2019) detailing the application of CNN for fault detection problems in civil infrastructures, electrical and mechanical machines, etc. The image processing aspect of 2D-CNN has been exploited by Cha *et al.* (2017) to detect cracks on the concrete surface and by Gulgec *et al.* (2019), to detect damage in steel gusset plate. Clearly, because of the photographic nature, these approaches are limited only to visible surface cracks and also conditioned on the visibility of the damage within the resolution of the captured image. (Abdeljaber *et al.* 2017) employed a 1D version of CNN to detect damage in a grandstand simulator's joints. This investigation processed the acceleration data collected from the vicinity of joints with CNN to classify them as either healthy or damaged. Abdeljaber *et al.* (2018) employed the 2D-CNN approach for damage detection in a benchmark structure introduced by the IASC-ASCE SHM task group. Their approach simultaneously records measurements from several sensors placed all over the structure with an aim to precisely localize the damage. (Sharma and Sen 2020b) validated numerically as well as experimentally the capability of the 1D-CNN approach in detection and localization of joint damage using measured member strain. Yet, in order to localize,

the application (also in Abdeljaber *et al.* (2017)) involved dedicated networks for each possible damage location, demanding exhaustive computation. Overall, through these studies, CNN has been identified as an efficient alternative to traditional model-based or shallow network-based SHM approaches while being sufficiently robust against uncertainties involved.

Prior to the damage classification with the CNN approach, this study decomposes the raw signal with the EMD technique. Pioneered by Huang *et al.* (1998), EMD has emerged as an efficient technique to decompose a raw signal into its component fundamental signals or IMFs. The term “intrinsic mode functions” signifies the hidden oscillations embedded within the time series (Dätig and Schlurmann 2004). The application of this approach for analysis of linear, non-linear, stationary and non-stationary signals is well appreciated in different research areas like process control (Srinivasan *et al.* 2007) medicine (Charleston-Villalobos *et al.* 2007), surface engineering (Zhang *et al.* 2008), system identification (Yang and Chang 2009) and speech recognition (Sharma *et al.* 2017) etc. Nevertheless, the application of EMD in SHM is relatively less explored.

Xu and Chen (2004) employed EMD to detect a sudden change of structural stiffness in a three-storey shear building. EMD was observed to be robust to external excitation in determining the instant and location of damage, even for the weaker and multiple damage cases. Yang and Chang (2009) proposed an indirect approach using EMD with Fast Fourier Transform for extracting bridge frequencies numerically as well as experimentally. Li *et al.* (2007) combined EMD with wavelet analysis to detect damage-induced anomalies in structural response. Yang and Chang (2009) proposed two methods using EMD and EMD-Hilbert transform for detection of damage instant, location and also pre-and post-damage modal parameters. (Rezaei and Taheri 2011) validated the efficacy of the EMD approach for damage estimation in a cantilever beam numerically and experimentally. Overall, the literature concludes EMD as a signal-based and model-free method for damage identification 216 requiring no prior knowledge of the structure.

### 3. Proposed EMD-2D-CNN

Two major expectations of SHM have been addressed in this article: temperature robustness and detection promptness. While for the detection promptness (and computational ease), the CNN environment can be an easy choice, to achieve the temperature robustness, a spatial-temporal correlation between responses at different sensor location can be exploited. This is due to the fact that while temperature affects the structure globally affecting all sensor responses equally, damage typically affects only a particular location. The spatial correlation between different sensor responses can therefore be exploited in order to detect damage locations more efficiently. An analogy of mode shape can be brought in this context which demonstrates the correlation between different sensor responses and less-likely to get affected by the ambient temperature variations. However, a damage can surely alter the undamaged mode shape. This aspect is equally true with the signals decomposed in to “simpler and well behaved” components (i.e., IMFs, analogous to modal response) and responses from several sensors are employed together for damage detection.

Eventually, this calls for the use of spatial dimension of responses as well, thereby justifying the need of a 2D version of CNN. Also, the detection network should be trained with temperature invariant data similar to fundamental mode-shapes that exhibit sufficient robustness against global temperature variance. However, typical frequency or time-frequency domain analysis approaches (FFT, wavelet analysis etc.) either bank on stationarity or requires a template signal to convolute.

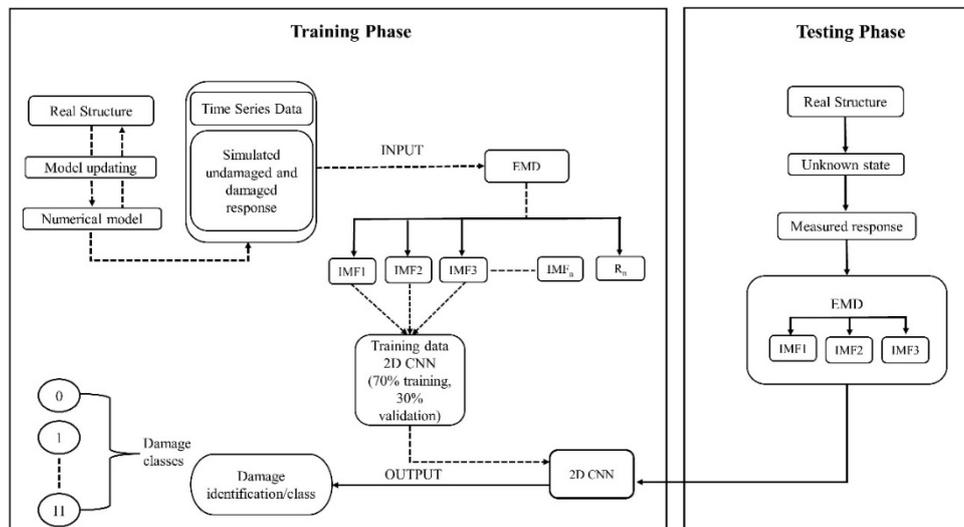


Fig. 2 The flowchart for the proposed algorithm

It's true that these approaches, by and large, are suitable for most condition assessment problems. However, the current state of the art demands the developed algorithms to suit even non-stationary signals aiming for real-time detection. The EMD approach, better suiting to this requirement, is therefore adopted in this study.

This article, therefore, combines the EMD with a 2D CNN network to identify damage from the ambient strain response under varying temperature conditions. In this approach, the noise-contaminated response data is firstly decomposed with the EMD into component IMFs from which strongly informative IMFs are chosen while neglecting all other less informative ones. This strategy takes basis on the assumption that the neglected IMFs constitute the signal's noise, which has no role in damage detection. Selection of the informative IMFs is undertaken through a correlation study following the works of Xun and Yan (2008), which are further classified against their corresponding damage cases by the 2D-CNN network. The proposed method is demonstrated in the following, along with the functioning of proposed approaches: i.e., EMD and CNN. A schematic of the proposed approach is presented in Fig. 2.

### 3.1 Empirical mode decomposition (EMD)

EMD self-adaptively decomposes a raw signal into component IMFs without taking the basis on any primary function (e.g., sinusoidal function in Fourier transform and mother wavelet in wavelet analysis) or preconceived filters (Flandrin *et al.* 2004). These IMFs are associated with energy at different time scales and contain essential characteristics of the real signal. The extraction of IMFs from the raw signal is achieved through an iterative process known as "sifting". Each IMF must satisfy the following two conditions (Huang *et al.* 1998):

- The total number of extrema and zero crossings should ideally be preserved or at most can differ by one.
- Mean of the envelopes defined by the local maxima and minima should ideally be zero.

For a time series,  $\mathbf{x}_{(1 \times t)} = \{\mathbf{x}_k; k = 1, \dots, t\}$ , with  $k$  being the time instants at which the signal is sampled, the extremes are firstly fitted through cubic splines in order to form the upper and lower envelopes. The envelope means,  $\mu(\mathbf{x})$ , is then subtracted from the original signal,  $\mathbf{x}$ , to obtain the first component,  $\mathbf{h}_{(1 \times t)}^1$ , of the sifting process.

$$\mathbf{h}^1 = \mathbf{x} - \mu(\mathbf{x}) \quad (1)$$

The resultant signal,  $\mathbf{h}^1$  is treated same as the original signal and the sifting process is repeated to obtain next mean subtracted component,  $\mathbf{h}^{11}$  as

$$\mathbf{h}^{11} = \mathbf{h}^1 - \mu(\mathbf{h}^1) \quad (2)$$

where  $\mu(\mathbf{h}^1)$  denotes the envelope mean of  $\mathbf{h}^1$ . The sifting process continues for  $k$  times until the  $\mathbf{h}^{1k}$  becomes the true IMF, as

$$\mathbf{h}^{1k} = \mathbf{h}^{1(k-1)} - \mu(\mathbf{h}^{1(k-1)}) \quad (3)$$

and is designated as the first IMF,  $\mathbf{i}_{(1 \times t)}^1$ . With each iteration, the sifting process produces a symmetrical signal with respect to mean zero. Huang *et al.* (1998) proposed criterion to stop the sifting process keeping the physical sense of the amplitude and frequency modulations into consideration. The criterion takes basis on the standard deviations of the deduced signal from the last two consequent sifting steps. Typically, the first IMF component,  $\mathbf{i}^1$ , is the component of the signal having the shortest period. The  $\mathbf{i}^1$  can then be separated from the original signal to obtain the residual,  $\mathbf{r}_{(1 \times t)}^1$  as

$$\mathbf{r}^1 = \mathbf{x} - \mathbf{i}^1 \quad (4)$$

$\mathbf{r}^1$ , containing the components of longer periods, is further treated as the raw signal to extract IMFs of longer periods. Recursion of this process leads to  $n$  empirical IMFs,  $\{\mathbf{i}^i, i = 1, \dots, n\}$  and a residue  $\mathbf{r}^n$  which can either be a mean trend or a constant. Ideally, the original signal  $\mathbf{x}$ , can be regained by summing up all the component IMFs and the residue as

$$\mathbf{x} = \sum_{i=1}^n \mathbf{i}^i + \mathbf{r}^n \quad (5)$$

It should be mentioned here that, in the context of infrastructural health monitoring, promptness is one of the most important aspect to achieve other than accuracy. The simpler computation process for EMD helps in making the overall process of damage detection faster. Typically, because of the advantages offered in terms of stability, convergence and second order smoothing, cubic spline interpolation is mostly approached for developing signal envelopes. Yet, this approach is susceptible to cause over/undershoot leading to mode aliasing and end effect. To circumvent these ill effects, Piece-wise Cubic Hermite Interpolation Polynomial (PCHIP) has been employed in this study supported by the works of Shulin *et al.* (2007). In several comparative studies (Rabbath and Corriveau 2019), PCHIP has been perceived as “shape-preserving” and “visually

pleasing" interpolation that can reconstruct the original signal much better than cubic spline interpolation and as such does not cause over/undershooting.

### 3.2 CNN

The article employs CNN to extract abstract features from the supplied IMF signals which are extracted from the response signals those are simultaneously recorded from multiple sensors. For a set of response time series  $\{\mathbf{x}_1 \cdots \mathbf{x}_{n_s}\}_{(t \times n_s)}$  of length  $t$  collected from  $n_s$  sensor locations, the IMFs are firstly extracted as  $\mathbf{i}_i^j$ , for,  $i = 1, \dots, n_s$  and  $j = 1, \dots, n_e$ . Here the notation  $\mathbf{i}_i^j$  denotes  $j^{\text{th}}$  IMF extracted from response measurement recorded from  $i^{\text{th}}$  sensor. Thus, with  $n_e$  numbers of IMFs taken into consideration, the overall input data that is supplied to CNN can be denoted as a three dimensional matrix  $\mathbf{I}_{(t \times n_s \times n_e)}$  from which each layer of size  $(t \times n_s)$  is fed to the CNN as individual image layers,  $I_{(t \times n_s)}$  for feature extraction purpose.

For each of these image layers,  $I_{(t \times n_s)} \in \mathbf{I}_{(t \times n_s \times n_e)}$ , CNN employs  $n_f$  local filters of dimension  $(m \times m)$ , i.e.,  $\{f_{k(m \times m)}; k = 1, \dots, n_f\}$ . These filters are convoluted on  $n^{\text{th}}$  window of the response  $I_{i:i+m-1, j:j+m-1(m \times m)}$  (denoted here on as  $I_{i,j,m}$  for the sake of compactness) to yield convoluted feature for  $n^{\text{th}}$  response window as  $c_{k,ijm}$  ( $n$  denotes the index of the window comprising the data  $I_{i,j,m}$  as

$$c_{\{k,ijm\}} = \phi(f_k \cdot I_{i,j,m} + b_k) \quad (6)$$

Thus,  $c_{\{k,ijm\}}$  can be regarded as the feature obtained through convoluting the filter  $f_k$  over the windowed signal data  $I_{i,j,m}$ . Here  $\langle s_1 \cdot s_2 \rangle$  denotes a convolutional operation between  $s_1$  and  $s_2$ .  $\phi(\cdot)$  denotes a nonlinear activation function and  $b_k$  denotes bias parameter. This process is followed for the entire signal  $I_{(t \times n_s)}$  by sliding this filter over the entire span (over the entire row and column space) of the data in order to obtain the feature map activated by the  $k^{\text{th}}$  filter as:  $\mathbf{C}_{k(q \times r)}$  where  $q = (t - m)/s + 1$  and  $r = (n_s - m)/s + 1$ . Here  $s$  is a hyper-parameter termed as stride that denotes the jump the filter makes between two successive convolutions.

In the following, pooling operation is performed over the feature map with an aim to dimensional reduction of the feature map as well as compress the information. Pooling techniques like max-pooling, min-pooling and average-pooling extracts max, min or average values from a window of the convoluted feature map of dimension  $(h \times h)$  with a stride value of  $S_h$ . Here  $h$  and  $S_h$  are hyper-parameters for the pooling operation and its value is generally decided through trial and error. The pooling window is slid over the convoluted feature map  $\mathbf{C}_j$  to down-sample the features to some manageable dimension. The pooled feature is denoted here as:  $\mathbf{P}_{k(\bar{q} \times \bar{r})}$  where and

$$\bar{q} = (q - h)/s_h + 1 \quad \text{and} \quad \bar{r} = (r - h)/s_h + 1.$$

$$\mathbf{P}_k = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,r} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{\bar{q},1} & p_{\bar{q},2} & \cdots & p_{\bar{q},\bar{r}} \end{bmatrix}_{(\bar{q} \times \bar{r})} \quad (7)$$

An arbitrary element in the  $\mathbf{P}_k$  matrix (say  $p_{(\bar{i}, \bar{j})}$ ) can be obtained as the maximum value

pooled from the block (i.e.,  $\mathbf{C}_{k(\bar{s}:\bar{s}+h,\bar{t}:\bar{t}+h)}$ ) of the convoluted matrix  $\mathbf{C}_k$  involving row space  $\{\bar{s}:\bar{s}+h\}$  and column space  $\{\bar{t}:\bar{t}+h\}$  where  $\bar{s} = (\bar{i} - 2) / h + s + 1$  and  $\bar{t} = (\bar{j} - 2) / h + s + 1$  can therefore be obtained as  $p_{(\bar{i},\bar{j})} = \max(\mathbf{C}_{k(\bar{s}:\bar{s}+h,\bar{t}:\bar{t}+h)})$ .

Thus, a two dimensional signal  $I$  of dimension  $(t \times n)_s$  is first convoluted to a feature map  $\mathbf{C}_K$  of dimension  $(q \times r)$  using a filter  $f_k$  of dimension  $(m \times m)$ . A pooling layer of dimension  $(h \times h)$  further down-samples the information to a matrix  $\mathbf{P}_k$  of dimension  $(\bar{q} \times \bar{r})$ . The dimensional reduction of information is performed several times by including several such stacked convolutional and pooling layers. Finally, after sufficient compression of the available information, the abstract feature map is classified using a fully connected layer and a softmax layer.

In the following, the proposed approach has been validated using numerical experiment. The further details of this approach are discussed specific to the problem accordingly.

#### 4. Numerical experiment

Current study attempts level 2 damage detection only (occurrence and location) coherent to the existing ML-based SHM researches (Jin *et al.* 2016, Weinstein *et al.* 2018, Zhu *et al.* 2019). With a primary objective of prompt detection of occurrence and location, severity detection is not attempted. However, the limiting level of damage that can be sensed with the proposed approach is investigated. In absence of real data corresponding to damaged and undamaged states of a structure, typically numerical experiments are adopted (Zhou *et al.* 2011, Yarnold and Moon 2015, Jin *et al.* 2016, Weinstein *et al.* 2018)s to validate the proposed approach. Nevertheless, even for the practical application of the proposed approach, damaged response of the real structure is required. While undamaged response collection is easy, the availability of damaged response of a physical structure under different damage conditions (location as well as severity) and varying temperature conditions is practically not possible. This study, therefore, resorts to a numerical model of a bridge structure, considered here in this study as the “test structure” (a proxy for the real structure) from which the response under different damage and temperature conditions are simulated. It should be mentioned here that, employment of synthetic damaged response instead of real damaged response is quite well-practiced approach in the literature (Jin *et al.* 2016, Weinstein *et al.* 2018, Zhu *et al.* 2019). Further details and modeling approach adopted for this “test structure” is demonstrated in the following. Accordingly, application of the proposed method for a real structure will follow these steps:

- (1) identification of the real structure in its undamaged condition,
- (2) creating its "Digital twin" (DT) to replicate the undamaged state,
- (3) simulation of damaged response from the DT and finally,
- (4) using this damaged response (along with real undamaged response) for training for the proposed supervised learning approach.

##### 4.1 The “test structure”

The “test structure” for this numerical experiment is basically a numerical model of a single span T-beam concrete bridge. The geometric details of this structure are presented in Table 1 while Fig. 3 presents a schematic representation of the bridge idealized as a beam-like structure. In practice, the bridges are often assumed to be fixed at one end while the other end to be supported

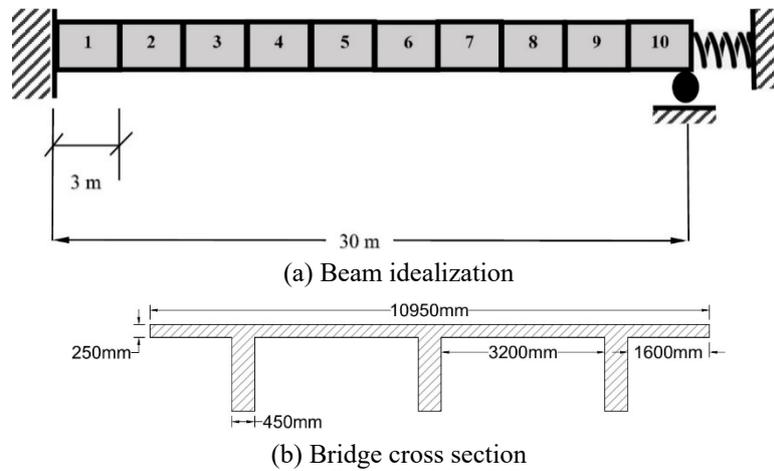


Fig. 3 Idealization of the numerical bridge model

Table 1 Assumed material and geometric properties for the numerical model

Material properties		Geometric and simulation details	
Density ( $kg/m^3$ )	2400	Area ( $m^2$ )	4.69
E (GPa)	26.5	$I_x$ ( $m^4$ )	2.79
Poisson's ratio	0.1	$I_y$ ( $m^4$ )	27.382
G (GPa)	12.086	$J$ ( $m^4$ )	30.18
Thermal coefficient ( $/^{\circ}C$ )	$1 \times 10^{-5}$	Temperature	$-30^{\circ}C$ to $60^{\circ}C$

by a roller. Nevertheless, compared to reality, this idealization of the support condition is only an over-simplification of the actual boundary condition. In reality, the rollers mostly don't allow complete release in the axial *dof*<sup>1</sup> and can be idealized as a finitely restrained support. This article attempts to model the supports more realistically in which the roller end is assumed to be finitely restrained by a linear spring. The schematic representation of the assumed boundary condition is presented in Fig. 3(a). This boundary condition allows finite thermal expansion, while the triggered thermal stresses can potentially alter the geometric stiffness of the bridge. The roller end is also considered to be rotationally restrained so that the bowing effect can be manifested. A much elaborate discussion on temperature and its effect on structural performance has been corroborated in (Sharma and Sen 2021a, b). The readers are requested to go through these articles for further details of the modeling.

#### 4.1.1 Modelling the effect of temperature

In order to replicate the impact of ambient temperature on real bridges, both, thermal expansion and bowing has been modeled along with their effects on geometric stiffness within a geometrically nonlinear Finite Element Model (FEM) of the bridge. An elaborate explanation of the modelling intricacies has been avoided here for the sake of compactness and for the relevant

<sup>1</sup>*dofs*: degrees of freedom

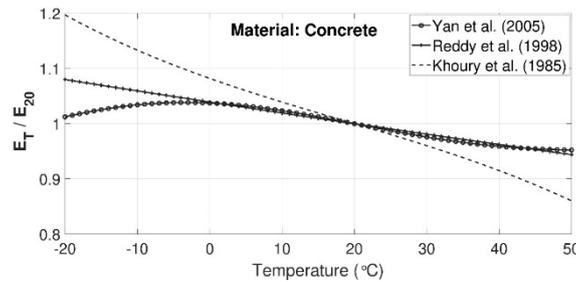
details, the readers may follow (Sharma and Sen 2021a, b). A brief detail of the same is, however, presented in the following. The girder bridge is subjected to varying temperatures ranging from  $-30^{\circ}\text{C}$  to  $60^{\circ}\text{C}$ . Apart from the uniform temperature change over the entire span, a thermal gradient of  $3^{\circ}\text{C}$  is also induced across the section to cause a thermal bowing.

It is perceived in general that temperature impacts stiffness more significantly compared to mass or damping. The property of structural material, being dependant on temperature, naturally impacts the material stiffness  $\mathbf{K}_m$ . The current study, therefore, employed temperature dependant material properties for elasticity, Poisson's ratio and thermal expansion coefficients for the component structural materials, i.e., concrete, and steel, following the studies of Khoury *et al.* (1985), Reddy and Chin (1998), Yan *et al.* (2005). The pertinent descriptions are presented in Figs. 4(a) and (b).

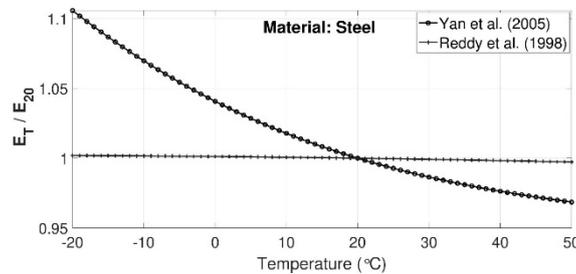
The temperature-induced expansion/contraction can cause supports to move which might cause a change in the support conditions, especially for the one assumed for this experiment. The linear spring at support ends can allow the expansion/contraction only finitely yielding axial thermal stress (tensile/ compressive) in addition to the allowed expansion. With a linear spring of stiffness  $k_l$  at support end, the developed axial stress ( $\sigma_a$ ) due to temperature variation  $\Delta T$  for a bridge of cross-section  $A$ , length  $l$ , equivalent material elasticity  $E$  and thermal expansion coefficient  $\alpha$  can be obtained as

$$\sigma_a = \frac{E \propto \Delta T}{1 + EA / k_t l} \quad (8)$$

Here  $\Delta T$  corresponds to a uniform temperature rise to  $T_{avg}$  from a bench-marked temperature  $T_0$  (temperature at no stress condition). With  $T_1$  and  $T_2$  being the temperatures of deck and the soffit of the bridge (commonly observed in bridges due to differential solar heating),



(a) Material: Concrete



(b) Material: Steel

Fig. 4 Effect of temperature on elastic modulus for concrete and steel

$T_{avg}$  is the average temperature of the cross-section (i.e.,  $(T_2 + T_1)/2$ ). Further, due to a linear gradient in temperature  $T_y (= |T_2 + T_1|/d)$  over the section depth  $d$ , thermal bowing may set in causing the beam-like structure to experience thermal-induced curvature. With support ends rotationally fixed, a tensile strain  $\varepsilon_\phi$  and associated tensile stress  $\sigma_b$  will eventually be developed, where

$$\varepsilon_\phi = 1 - \frac{\sin(l\phi/2)}{l\phi/2}; \quad \sigma_b = E\varepsilon_\phi \quad (9)$$

Here,  $\phi$  is the radius of curvature due to bowing calculated as  $\phi = \alpha T_y$ . It should be noted that a gradient in temperature only in the vertical direction is considered while it has been assumed that the temperature along the bridge axis and across the deck is constant. Evidently, two stresses are generated in the structure due to a change in temperature which can potentially impact the overall stiffness through altering the geometric stiffness  $\mathbf{K}_g$ . With  $\sigma_t$  being the effective stress (i.e.,  $\sigma_t = \sigma_a - \sigma_b$ ), the corresponding geometric stiffness matrix  $\mathbf{K}_g$  can be obtained following (Gavin 2012) and further can be added to the material stiffness matrix  $\mathbf{K}_m$ .

#### 4.1.2 FEM of the “test structure”

The bridge is numerically simulated using the geometrically nonlinear FEM approach in which the entire span is discretized into ten segments of equal length. Each of these segments is modeled using a two-noded three-dimensional Euler-Bernoulli beam element involving geometric nonlinearity and its analytical consequence, i.e., geometric stiffness. The geometric stiffness accounts for the effect of finite deformation on the structural stiffness due to the induced thermal prestress. Each node of the employed 3D beam elements is defined with six *dof*s (three translational and three rotational). The material properties are defined with their dependence on temperature following Fig. 4. 2% Rayleigh damping has been assumed for simulation.

Various damage scenarios are introduced randomly in terms of numerical reduction in material elasticity ( $\mathbf{E}$ ) in different segments under temperature variability. In this paper, the assumed damage severity is considered to be varying within the range of 20-80% of the original material elasticity, which is further categorized into three levels: weak (20-40%), moderate (40-60%), and strong (60-80%) in order to investigate the efficacy of the proposed algorithm in detecting damages of varying severities. It should be noted that the mentioned damage severities are given in terms of elasticity loss and not to be confused with the actual reduction in elasticity which is a result of combined effect of elasticity loss due to damage and temperature.

To simulate response from the “test structure”, it is excited in its all vertical *dofs* with a stationary white Gaussian noise (SWGN) force model of distribution  $N(0, \mathbf{Q})^2$  considered as the ambient forcing acting on the structure. Each of the segments is assumed to be instrumented at its respective middle point with a strain gauge patched at its bottom surface (soffit of the T-girder) recording axial strain component along the axis of the bridge. The dynamic strain is sampled from these gauges at a constant sampling frequency of 50 Hz for 1024 long time-series data. Each simulation is associated to one realization of damaged segment ( $\epsilon \in \{0, 1, \dots, 10\}$ , 0 denotes undamaged), severity (in %  $\epsilon \in \{20, 80\}$  of material elasticity) and ambient temperature ( $\epsilon \in \{-30^\circ\text{C}, 60^\circ\text{C}\}$ ) drawn from uniform distributions (discrete or continuous):  $U^d([0; 10])$ ,

<sup>2</sup> A distribution  $N(0, \mathbf{Q})^2$  signifies a zero mean SWGN of variance  $\mathbf{Q}$

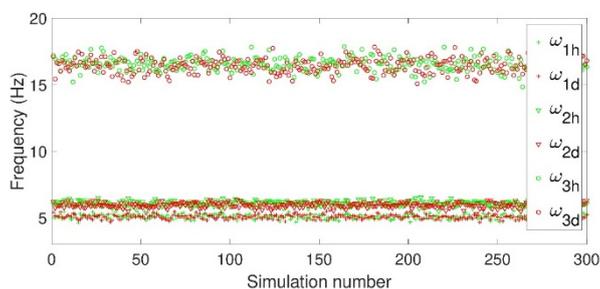


Fig. 5 Variation of healthy and damaged frequencies under varying temperature

$U^c([20; 80])$   $U^c([-30; 60])$  respectively <sup>3</sup>.

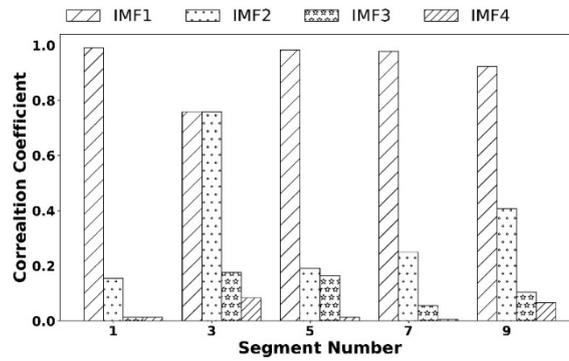
Further, to replicate the real field scenarios, the simulated responses are contaminated with various levels of measurement noise. The added noise is assumed to be following an SWGN model and defined with *signal-to-noise* (*snr*) measure in which a  $x\%$  *snr* signifies that the ratio of standard deviations of noise to original uncontaminated signal is  $x\%$ . For each noise level, a set of 1100 cases are simulated for 11 damage scenarios (10 damaged + 1 undamaged) under 100 instances of ambient SWGN forcing. Each of these simulation cases yielded a strain data of dimension of  $1024 \times 10$  attributed to strain responses from 10 members of 1024 data points in time. A sample comparison is presented between healthy and damaged frequencies under temperature variability in Fig. 5 that demonstrates the substantial impact of temperature on the frequencies.

#### 4.2 IMF extraction and training

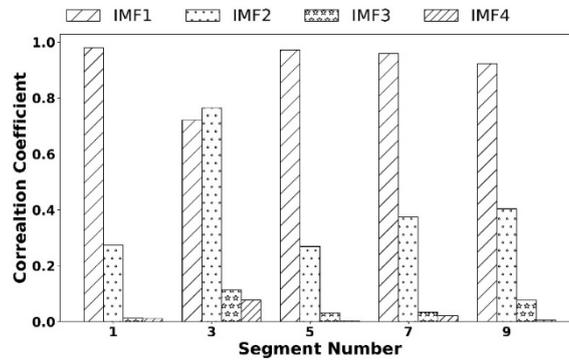
The contaminated strain responses are further put through the EMD technique to extract the pertinent IMFs sensitive to the presence of damage in the structure. In this regard, a correlation study is undertaken for both damaged and undamaged responses in which the correlation between IMFs and the original signal is estimated. Following the works of Xun and Yan (2008), the adopted strategy assumes that the poorly correlated IMFs constitute the less significant portion of the signal majorly originated due to the noise present in the signal. A sample of this study is presented in Fig. 6. The first two IMFs are consistently perceived as the IMFs of strong correlation and therefore should always be considered in the subsequent training steps. For signals with high levels of noise contamination, the third IMF is also found to be moderately correlated at times. The trend is also observed to be uniform for both damaged and undamaged responses

While the inclusion of more numbers of IMFs in the calculation increases the computation exponentially, it does not necessarily increase the precision or accuracy in detection substantially. A separate study has been taken up that involves increasing numbers of IMFs in the CNN-based detection network to check how the level of detection accuracies varies with the number of IMFs considered. Four such tests are undertaken with 1/2/3 and 5 IMFs extracted from the same signal and the detection accuracies are monitored. The results are presented in Fig. 7(b) where from it is evident that up to the third IMF the accuracy levels are increasing. Any further increase in IMF does not benefit in terms of accuracy while making the approach compute intensive. An example

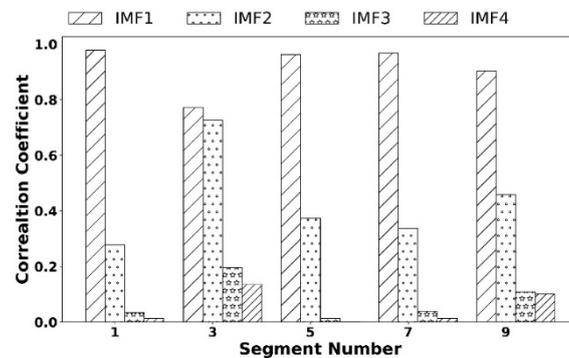
<sup>3</sup>  $U^{c/d}([a; b])$  denotes a uniform distribution bounded within the closed set  $[a; b]$ . Superscript  $c$  and  $d$  signifies continuous or discrete type distribution



(a) Correlation coefficients for undamaged case



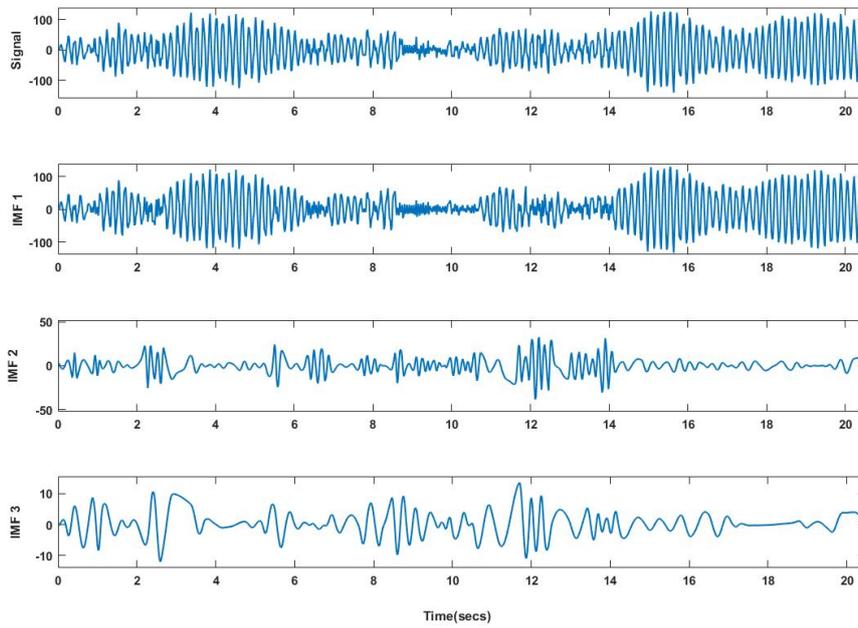
(b) Correlation coefficients for damaged case



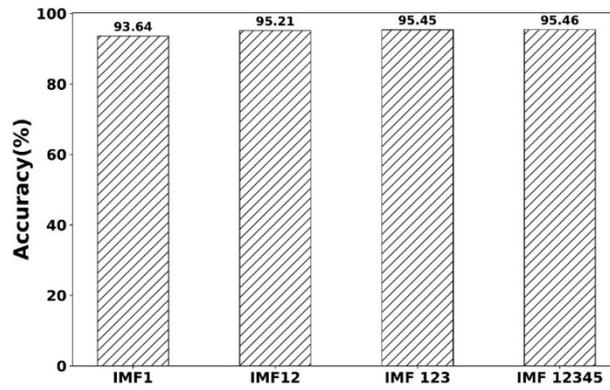
(c) Correlation coefficients with 10% snr

Fig. 6 The correlation coefficients of first four IMFs with signal at 20°C

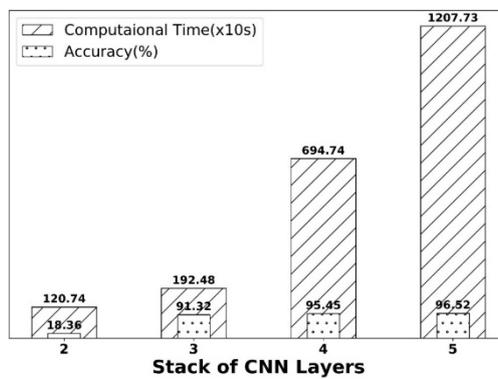
of corresponding IMFs are presented in Fig. 7(a). Thus, a maximum of three IMFs is chosen for this study that is manageable within computational limitation while being sufficient enough to deal with signals with any practical levels of noise contamination. Eventually, the EMD approach decomposed a 2D strain data of dimension  $(1024 \times 10)$  into 3D IMF data of dimension  $(1024 \times 10 \times 3)$ . In the following, the IMF data is fed to the proposed 2D CNN network in its input layer while the corresponding damage cases are fed in the output layer. 70% of the simulated data is used for training and remaining 30% for validation and testing. Each damage identification index



(a) The first three IMFs of the simulated



(b) Effect of increasing number of IMFs on accuracy levels



(c) Effect of increasing stack of convolutional and pooling layers

Fig. 7 Performance analysis of the proposed algorithm

is one component from a categorical vector  $C$  of dimension  $(11 \times 1)$ . Categorical vector  $C$  stores the eleven possible damage classes (one healthy and 10 damaged). Consideration of multiple damage cases will eventually call for more damage cases jeopardizing the promptness aspect. It should be noted here that, the proposed approach can be categorized as a near-real-time approach detecting the damage instantly once the 1024 long time series is available. Accordingly, for pragmatic reasons, only single damage cases have been considered presuming only one damage is expected to occur within the time frame of sampled measurements. For real-life SHM problems, it is highly unlikely to have two disjointed spans getting damaged at once. Of course, for such less-likely scenarios, damage occurrence can still be detected without a perfect localization.

The network design needs certain hyper-parameters to be decided, e.g., filter size and numbers, number of stacking layers and sampling windows etc., which is usually undertaken by recursive trials with an aim to improve detection accuracy with the training dataset (Abdeljaber *et al.* 2017, Avci *et al.* 2017). In this process, it is not uncommon to have parameters more than the length of the signal. While, typical MNIST problem (handwritten digit recognition) requires only two stacks of CNN, keeping the complexity of the detection problem into consideration, current study investigates the required numbers of stacking that can yield good accuracy with the training data. The size of filters has also been optimized in this attempt in order to not miss the important information while being within manageable computation demand. In such contexts, ancillary algorithms are typically employed, especially for big data problems, to remove weaker nodes. Nothing such was however attempted in this study since the computation was already within acceptable limits.

The selection of the required layers of convolution and pooling in the employed CNN network is therefore an iterative effort. In this article, a rigorous numerical study is undertaken involving different stacks of CNN layers (2/3/4/5) and corresponding accuracy levels with the training data are monitored. The results are presented in Fig. 7(c). It has been found that four stacks of convolution and pooling are sufficient to yield the required accuracy even with the most contaminated signal, while a further increase is observed to only exhausting the computational resource at no extra benefits (cf. Fig. 7(c)). A four-layer 2D CNN network is therefore standardized for all numerical case studies undertaken in this study.

The finally standardized CNN architecture is presented in Table 2 with details of all the layers employed. Overall, this 2D-CNN network constitutes four stacks of convolution and pooling layers apart from regular batch normalization, dropout, and activation layers (relu layer). The input layer consumes a 3D image-like input of dimension  $1024 \times 10 \times 3$  (similar to RGB matrix of image inputs) and the output layer assigns one of the eleven possible damage cases to this input. A learning rate of 0.001 and Stochastic Gradient Descent with momentum (SGDM) is used as an optimiser with a batch size of 50. The network is trained over the 300 number of epochs. The final feature map is classified using a fully connected layer with the help of the softmax activation function. Table 2 lists all the component layers elaborately along with the relevant dimensions.

The network is first tested for its efficiency in detecting damages in the “test structure” under a moderate damage level (40-60%) and under a moderate noise contamination level (2% *snr*). The network is trained using strain responses corresponding to 1100 test cases (11 damage cases  $\times$  100 simulations under random damage level, location, ambient temperatures and forcing). The proposed 2D-CNN network is observed to be very precise in detecting the occurrence of damages while pinpointing their locations for most cases. The algorithm is also found to be causing significantly low numbers of false alarms. The detection accuracy in terms of True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) are presented in Table 3. The

Table 2 Details of CNN architecture

Layer	Name	Operation	Dimensions
1	'image input'	Image Input	$1024 \times 10 \times 3$ images
2	'conv_1'	Convolution	512, $3 \times 3 \times 3$ convolutions with stride [1 1] and padding [1 1 1]
3	'batchnorm_1'	Batch Normalization	Batch normalization with 512 channels
4	'relu_1'	ReLU	Activation function
5	'maxpool_1'	Max Pooling	$2 \times 2$ max pooling with stride [2 2]
6	'conv_2'	Convolution	128, $3 \times 3 \times 512$ convolutions with stride [1 1] and padding [1 1 1]
7	'batchnorm_2'	Batch Normalization	Batch normalization with 128 channels
8	'relu_2'	ReLU	Activation function
9	'maxpool_2'	Max Pooling	$2 \times 2$ max pooling with stride [2 2]
10	'conv_3'	Convolution	64, $3 \times 3 \times 128$ convolutions with stride [1 1] and padding [1 1 1]
11	'batchnorm_3'	Batch Normalization	Batch normalization with 64 channels
12	'relu_3'	ReLU	Activation function
13	'maxpool_3'	Max Pooling	$2 \times 2$ max pooling with stride [2 2]
14	'conv_4'	Convolution	16, $3 \times 3 \times 64$ convolutions with stride [1 1] and padding [1 1 1]
15	'batchnorm_4'	Batch Normalization	Batch normalization with 16 channels
16	'relu_4'	ReLU	Activation function
17	'maxpool_4'	Max Pooling	$2 \times 2$ max pooling with stride [2 2]
18	'dropout'	Dropout	50 % dropout
19	'fc'	Fully Connected	11 fully connected layer
20	'softmax'	Softmax	softmax
21	'classoutput'	Classification Output	'0' and 10 other classes

relevant numerical expressions are also given therein. Additionally, the ratio of TP and TN, Accuracy (ACC), precision (PPV) and Youden's index (YI) are also presented for a better judgment of the efficacy of the proposed algorithm. YI or Informedness index is single statistic that assesses a multi-class classifier's performance while giving false positive and negative detection equal importance. YI ranges from 0 to 1, where 0 and 1 signifies a very poor and perfect performance classifier respectively. In contrast to a random guess, YI facilitates with an informed decision that takes all predictions in to consideration.

Next, the noise sensitivity of the proposed algorithm with the defined network is investigated. SWGN noise of four different strengths (0/1/2/5 % *snr*) are experimented with. It has been observed that the proposed network can efficiently identify the damage locations even in presence of noise as high as 5% *snr*. For a particular level of induced damage, the level of accuracy is, however, observed to be decreasing with the increasing levels of noise (cf. Table 4). To have a conclusive remark on the noise sensitivity of the proposed algorithm and the levels of damage that can be practically identified with this proposed algorithm, a detailed study is undertaken

Table 3 Performance assessment for a moderate damage case with 2% noise level

	TP	TN	FP	FN	TPR	TNR	ACC	PPV	YI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Class	Numerical definitions								
	$n_{1 1}^d$	$n_{0 0}^d$	$n_{1 0}^d$	$n_{0 1}^d$	$\frac{(1)}{(1) + (4)}$	$\frac{(2)}{(2) + (3)}$	$\frac{(1) + (2)}{(1) + (2) + (3) + (4)}$	$\frac{(1)}{(1) + (4)}$	$(5) + (6) - 1$
0	81	966	34	19	0.81	0.96	0.95	0.70	0.77
1	100	948	0	0	1.00	1.00	1.00	1.00	1.00
2	100	948	0	0	1.00	1.00	1.00	1.00	1.00
3	99	947	0	1	0.99	1.00	0.99	1.00	0.98
4	100	948	0	0	1.00	1.00	1.00	1.00	1.00
5	100	948	0	0	1.00	1.00	1.00	1.00	1.00
6	100	948	0	0	1.00	1.00	1.00	1.00	1.00
7	100	948	0	0	1.00	1.00	1.00	1.00	1.00
8	100	948	0	0	1.00	1.00	1.00	1.00	1.00
9	100	947	1	0	1.00	0.99	0.99	0.99	0.99
10	66	981	19	34	0.66	0.98	0.95	0.77	0.64

Notation  $n_{a|b}^d$  counts the instances of network decision for segment d as a while the truth is b.

Notation  $n_c^d$  counts the instances for segment d with state c.

a, b, c are Boolean variables taking values 0 or 1 denoting healthy and damaged states respectively.

Table 4 Detection efficiency of the proposed algorithm for different damage and noise contamination levels

Damage severity	Weak: (20-40%)			
Noise (%snr)	0	1	2	5
Time (secs)	10869	11410	11334	11470
Accuracy (%)	86.36	79.90	79.70	77.88
Damage severity	Moderate: (40-60%)			
Noise (%snr)	0	1	2	5
Time (secs)	6600	6947	6958	7020
Accuracy (%)	95.90	95.45	93.03	91.80
Damage severity	Strong: (60-80%)			
Noise (%snr)	0	1	2	5
Time (secs)	5317	5297	5863	5880
Accuracy (%)	100.00	100.00	100.00	100.00

involving different noise levels and damage severities.

As mentioned previously in this article, the possible damage scenarios are segmented into three severity levels (weak: 20-40%, moderate: 40-60%, and strong: 60-80%). It has been presumed that the employment of SHM for damages over 80% is not pragmatic, while any damage of severity

< 20% is also not feasible. The latter assumption has further been confirmed using a separate study which showed very poor performance with the algorithm even with moderate noise contamination levels (2-5%). The reason for this poor performance can be attributed to the imperceptible change induced by such weak damage (1.96% change in the first frequency for a 10% damage in the second segment). Thus, this study focused on detecting practical levels of damage within the range (20-80%). However, for weaker damage cases, an alternate approach has been suggested later in this article.

Three sets of experiments are undertaken for the three mentioned damage severities for all four predefined noise levels. It has been found that estimation accuracy depends on noise sensitivity and damage levels. For uncontaminated signals, all of the damage levels are estimated without much difference in the accuracies. However, for noisy signals, damage levels are found to play a major role in the detection accuracies. On the other hand, signals with strong damage severities are found to be sufficiently robust against noise contamination levels and therefore, the accuracy levels are mostly undeterred for all noise levels. Nevertheless, signals of weak damage severities are found to be significantly sensitive to noise.

Overall, the algorithm has been found to be sufficiently robust for all damage severities for moderate noise contamination levels. The signals, perceived to be contaminated with high levels of noise ( $> 5\%$  *snr*), or from a weakly damaged ( $< 20\%$ ) structure is advised to be taken special care in order to ensure certainty in detection. A Probability of Detection (POD) measure can be an efficient approach in this endeavor. For this, several signals from the same structure should be put through the proposed algorithm and the detection decisions have to be judged based on the frequency of the decisions provided by the algorithm. Mostly, the segment with actual damage should ideally be having maximum POD measure. For an example case study with 10% damage, the POD measures are presented in Fig. 8 where it can be observed that the proposed algorithm can still detect the damage location probabilistically. The precision of the detection will, however, depend on the noise levels and damage severities.

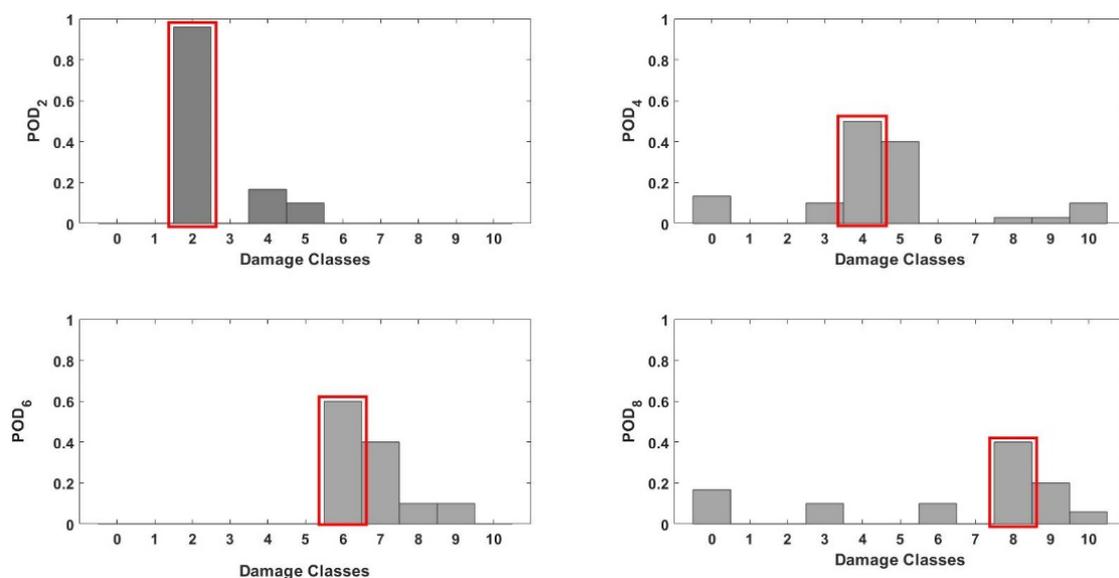


Fig. 8 Probability of Detection for < 20% damage case

## 5. Conclusions

A deep learning-based 2D-CNN approach coupled with EMD is adopted in this study to develop a near-real-time damage detection algorithm that is robust against ambient thermal uncertainty and practical levels of measurement noise. The EMD approach decomposes the 2D array of dynamic strain signals, recorded from a network of strain gauges, into damage-sensitive IMFs. Further, abstract spatio-temporal features embedded within these IMFs, sensitive only to damage, are extracted automatically by the 2D-CNN approach. The adopted CNN approach does not require a preconceived feature as the basis for damage identification. Instead, it extracts the features all by itself from the signal and categorizes them into corresponding damage cases. The selection of temperature invariant yet damage-sensitive features is thus automated, circumventing any forms of human intervention and associated bias. Overall, the proposed approach reduces the impact of sensor noise contamination while making the damage detection approach feature-free.

Considering the non-availability of structural response data in its healthy and damaged conditions under varying temperature conditions, the present study resorted to numerical experiments (a FEM as a proxy for real structure) to validate the proposed ML-based SHM approach. For this, a bridge structure modeled using a geometrically nonlinear FEM involving thermal effects in its material and geometric stiffness. From this model, the impact of ambient temperature on structural stiffness is perceived to be significant compared to the structural damage.

FEM simulated responses are subsequently analyzed with the proposed approach in which the selection of the IMFs is found to play a major role in deciding the detection accuracy. Supported by a correlation study following (Xun and Yan 2008), the first three IMFs are found to be sensitive as well as affecting the accuracies positively. The inclusion of any further IMFs was either redundant or detrimental from the perspective of the computational economy.

The further elaborate validation study established that the proposed algorithm is very accurate and precise in damage identification as long as the damage levels are  $\geq 20\%$  with noise  $\leq 5\%$ . The coupling of EMD with 2D CNN approach exploited the spatio-temporal correlations in the recorded signal arrays which is much less sensitive to ambient thermal variability. The approach ensured better precision even with moderate to high noise-contaminated signals. For damage level  $< 20\%$  and noise levels  $> 5\%$ , a POD approach is proposed that takes a frequentist approach in search of the most probable damage location.

## Acknowledgments

This study was funded by Aeronautics Research & Development Board (DRDO), New Delhi, India through grant file no. ARDB/01/1051907/M/I.

## References

- Abdel-Hamid, O., Mohamed, A.R., Jiang, H. and Penn, G. (2012), "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition", *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, March, pp. 4277-4280. <https://doi.org/10.1109/ICASSP.2012.6288864>
- 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>
- Abdeljaber, O., Avci, O., Kiranyaz, M.S., Boashash, B., Sodano, H. and Inman, D.J. (2018), "1-D CNNs for structural damage detection: Verification on a structural health monitoring benchmark data", *Neurocomputing*, **275**, 1308-1317. <https://doi.org/10.1016/j.neucom.2017.09.069>
- Alampalli, S. (2000), "Effects of testing, analysis, damage, and environment on modal parameters", *Mech. Syst. Signal Process.*, **14**(1), 63-74. <https://doi.org/10.1006/mssp.1999.1271>
- Avci, O., Abdeljaber, O., Kiranyaz, S. and Inman, D. (2017), "Structural damage detection in real time: Implementation of 1D convolutional neural networks for SHM applications", *Proceedings of the Society for Experimental Mechanics Series*, Volume 7, pp. 49-54. [https://doi.org/10.1007/978-3-319-54109-9\\_6](https://doi.org/10.1007/978-3-319-54109-9_6)
- Cha, Y.J., Choi, W. and Büyüköztürk, O. (2017), "Deep learning-based crack damage detection using convolutional neural networks", *Comput.-Aided Civil Infrastr. Eng.*, **32**(5), 361-378. <https://doi.org/10.1111/mice.12263>
- Charleston-Villalobos, S., González-Camarena, R., Chi-Lem, G. and Aljama-Corrales, T. (2007), "Crackle sounds analysis by empirical mode decomposition: Nonlinear and nonstationary signal analysis for distinction of crackles in lung sounds", *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine & Biology Society*, **26**(1), pp. 40-47.
- Cornwell, P., Farrar, C.R., Doebling, S.W. and Sohn, H. (1999), "Environmental variability of modal properties", *Experim. Techniq.*, **23**(6), 45-48. <https://doi.org/10.1111/j.1747-1567.1999.tb01320.x>
- LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W. and Jackel, L. (1989a), "Handwritten digit recognition with a back-propagation network", *Adv. Neural Inform. Process. Syst.*, **2**.
- Le Cun, Y., Jackel, L.D., Boser, B., Denker, J.S., Graf, H.P., Guyon, I., Henderson, D., Howard, R.E. and Hubbard, W. (1989b), "Handwritten digit recognition: Applications of neural network chips and automatic learning", *IEEE Commun. Magazine*, **27**(11), 41-46. <https://doi.org/10.1109/35.41400>
- Dätig, M. and Schlurmann, T. (2004), "Performance and limitations of the Hilbert-Huang transformation (HHT) with an application to irregular water waves", *Ocean Eng.*, **31**(14-15), 1783-1834. <https://doi.org/10.1016/j.oceaneng.2004.03.007>
- De Oliveira, M.A., Araujo, N.V., Da Silva, R.N., Da Silva, T.I. and Epaarachchi, J. (2018), "Use of Savitzky-Golay filter for performances improvement of SHM systems based on neural networks and distributed PZT sensors", *Sensors*, **18**(1), 152. <https://doi.org/10.3390/s18010152>
- dos Santos, F.L.M., Peeters, B., Lau, J., Desmet, W. and Goes, L.C.S. (2015), "The use of strain gauges in vibration-based damage detection The use of strain gauges in vibration-based damage detection", *J. Physics: Conference Series*, **628**(1), 012119. <https://doi.org/10.1088/1742-6596/628/1/012119>
- Duque, A.B., Santos, L.L.J., Macêdo, D. and Zanchettin, C. (2019), "Squeezed very deep convolutional neural networks for text classification", *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 11727, pp. 193-207. [https://doi.org/10.1007/978-3-030-30487-4\\_16](https://doi.org/10.1007/978-3-030-30487-4_16)
- Flandrin, P., Rilling, G. and Gonçalves, P. (2004), "Empirical mode decomposition as a filter bank", *IEEE Signal Process. Lett.*, **11**(2), 112-114. <https://doi.org/10.1109/LSP.2003.821662>
- Gavin, H. (2012), "Geometric stiffness effects in 2D and 3D frames", Department of Civil and Environment Engineering, Edmund T. Pratt School of Engineering, Duke University, pp. 1-14.
- Glisic, B., Chen, J. and Hubbell, D. (2011), "Streicker Bridge: a comparison between Bragg-grating long-gauge strain and temperature sensors and Brillouin scattering-based distributed strain and temperature sensors", In: *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2011*, Volume 7981, p. 79812C. <https://doi.org/10.1117/12.881818>
- Gulgec, N.S., Takáč, M. and Pakzad, S.N. (2019), "Convolutional neural network approach for robust structural damage detection and localization", *J. Comput. Civil Eng.*, **33**(3), p. 04019005. [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000820](https://doi.org/10.1061/(asce)cp.1943-5487.0000820)
- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H. (1998), "The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **454**(1971), pp. 903-995. <https://doi.org/10.1098/rspa.1998.0193>

- Ince, T., Kiranyaz, S., Eren, L., Askar, M. and Gabbouj, M. (2016), "Real-time motor fault detection by 1-D convolutional neural networks", *IEEE Transact. Indust. Electron.*, **63**(11), 7067-7075. <https://doi.org/10.1109/TIE.2016.2582729>
- Jin, C., Jang, S., Sun, X., Li, J. and Christenson, R. (2016), "Damage detection of a highway bridge under severe temperature changes using extended Kalman filter trained neural network", *J. Civil Struct. Health Monitor.*, **6**(3), 545-560. <https://doi.org/10.1007/s13349-016-0173-8>
- Khoury, G.A., Grainger, B.N. and Sullivan, J.E. (1985), "Transient thermal strain of concrete: literature review, conditions within specimen and behaviour of individual constituents", *Magaz. Concrete Res.*, **37**(132), pp. 131-144. <https://doi.org/10.1680/mac.1985.37.132.131>
- Kiranyaz, S., Ince, T. and Gabbouj, M. (2016), "Real-time patient-specific ECG classification by 1-D convolutional neural networks", *IEEE Transact. Biomed. Eng.*, **63**(3), 664-675. <https://doi.org/10.1109/TBME.2015.2468589>
- Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012), "Imagenet classification with deep convolutional neural networks", *Adv. Neural Inform. Process. Syst.*, **25**, 1097-1105.
- Kromanis, R., Kripakaran, P. and Harvey, B. (2016), "Long-term structural health monitoring of the Cleddau bridge: Evaluation of quasi-static temperature effects on bearing movements", *Struct. Infrastr. Eng.*, **12**(10), 1342-1355. <https://doi.org/10.1080/15732479.2015.1117113>
- Kullaa, J. (2009), "Eliminating environmental or operational influences in structural health monitoring using the missing data analysis", *J. Intell. Mater. Syst. Struct.*, **20**(11), 1381-1390. <https://doi.org/10.1177/1045389X08096050>
- Le Cun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W. and Jackel, L. (1989), "Handwritten digit recognition with a back-propagation network", *Adv. Neural Inform. Process. Syst.*, **2**.
- Li, H.L., Deng, X. and Dai, H. (2007), "Structural damage detection using the combination method of EMD and wavelet analysis", *Mech. Syst. Signal Process.*, **21**(1), 298-306. <https://doi.org/10.1016/j.ymsp.2006.05.001>
- Rabbath, C.A. and Corriveau, D. (2019), "A comparison of piecewise cubic Hermite interpolating polynomials, cubic splines and piecewise linear functions for the approximation of projectile aerodynamics", *Defence Technology*, **15**(5), 741-757. <https://doi.org/10.1016/j.dt.2019.07.016>
- Reddy, J.N. and Chin, C.D. (1998), "Thermomechanical analysis of functionally graded cylinders and plates", *J. Thermal Stresses*, **21**(6), 593-626. <https://doi.org/10.1080/01495739808956165>
- Rezaei, D. and Taheri, F. (2011), *Structural Health Monitoring*, John Wiley & Sons. <https://doi.org/10.1177/1475921710373298>
- Sharma, S. and Sen, S. (2020), "One-dimensional convolutional neural network-based damage detection in structural joints", *J. Civil Struct. Health Monitor.*, **10**(5), 1057-1072. <https://doi.org/10.1007/s13349-020-00434-z>
- Sharma, S. and Sen, S. (2021a), "Bridge Damage Detection in Presence of Varying Temperature Using Two-Step Neural Network Approach", *J. Bridge Eng.*, **26**(6), 04021027. [https://doi.org/10.1061/\(asce\)be.1943-5592.0001708](https://doi.org/10.1061/(asce)be.1943-5592.0001708)
- Sharma, S. and Sen, S. (2021b), "Damage Detection in Presence of Varying Temperature Using Mode Shape and a Two-Step Neural Network", In: *Recent Advances in Computational Mechanics and Simulations*, pp. 285-299. [https://doi.org/10.1007/978-981-15-8138-0\\_23](https://doi.org/10.1007/978-981-15-8138-0_23)
- Sharma, R., Vignolo, L., Schlotthauer, G., Colominas, M.A., Rufiner, H.L. and Prasanna, S.R.M. (2017), "Empirical mode decomposition for adaptive AM-FM analysis of speech: A review", *Speech Communication*, **88**, 39-64. <https://doi.org/10.1016/j.specom.2016.12.004>
- Shulin, L., Haifeng, Z., Hui, W. and Rui, M. (2007), "Application of improved EMD algorithm for the fault diagnosis of reciprocating pump valves with spring failure", In: *2007 9th International Symposium on Signal Processing and Its Applications*, pp. 1-4. <https://doi.org/10.1109/ISSPA.2007.4555473>
- Sohn, H., Farrar, C.R., Hemez, F.M., Shunk, D.D., Stinernes, D.W., Nadler, B.R. and Czarnecki, J.J. (2003), "A review of structural health monitoring literature: 1996-2001", Los Alamos National Laboratory, USA, 1.
- Srinivasan, R., Rengaswamy, R. and Miller, R. (2007), "A modified empirical mode decomposition (EMD)

- process for oscillation characterization in control loops”, *Control Eng. Practice*, **15**(9), 1135-1148. <https://doi.org/10.1016/j.conengprac.2007.01.014>
- Sun, L., Shang, Z., Xia, Y., Bhowmick, S. and Nagarajaiah, S. (2020), “Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection”, *J. Struct. Eng.*, **146**(5), 04020073. [https://doi.org/10.1061/\(asce\)st.1943-541x.0002535](https://doi.org/10.1061/(asce)st.1943-541x.0002535)
- Weinstein, J.C., Sanayei, M. and Brenner, B.R. (2018), “Bridge damage identification using artificial neural networks”, *J. Bridge Eng.*, **23**(11), 04018084. [https://doi.org/10.1061/\(asce\)be.1943-5592.0001302](https://doi.org/10.1061/(asce)be.1943-5592.0001302)
- Xia, Q., Cheng, Y., Zhang, J. and Zhu, F. (2017), “In-service condition assessment of a long-span suspension bridge using temperature-induced strain data”, *J. Bridge Eng.*, **22**(3), 04016124. [https://doi.org/10.1061/\(asce\)be.1943-5592.0001003](https://doi.org/10.1061/(asce)be.1943-5592.0001003)
- Xu, Y.L. and Chen, J. (2004), “Structural damage detection using empirical mode decomposition: experimental investigation”, *J. Eng. Mech.*, **130**(11), 1279-1288. [https://doi.org/10.1061/\(asce\)0733-9399\(2004\)130:11\(1279\)](https://doi.org/10.1061/(asce)0733-9399(2004)130:11(1279))
- Xun, J. and Yan, S. (2008), “A revised Hilbert-Huang transformation based on the neural networks and its application in vibration signal analysis of a deployable structure”, *Mech. Syst. Signal Process.*, **22**(7), 1705-1723. <https://doi.org/10.1016/j.ymsp.2008.02.008>
- Yan, A.M., Kerschen, G., De Boe, P. and Golinval, J.C. (2005), “Structural damage diagnosis under varying environmental conditions - Part I: A linear analysis”, *Mech. Syst. Signal Process.*, **19**(4), 847-864. <https://doi.org/10.1016/j.ymsp.2004.12.002>
- Yang, Y.B. and Chang, K.C. (2009), “Extraction of bridge frequencies from the dynamic response of a passing vehicle enhanced by the EMD technique”, *J. Sound Vib.*, **322**(4-5), 718-739. <https://doi.org/10.1016/j.jsv.2008.11.028>
- Yarnold, M.T. and Moon, F.L. (2015), “Temperature-based structural health monitoring baseline for long-span bridges”, *Eng. Struct.*, **86**, 157-167. <https://doi.org/10.1016/j.engstruct.2014.12.042>
- Zhang, H., Qi, X., Sun, X. and Fan, S. (2008), “Application of Hilbert-Huang transform to extract arrival time of ultrasonic lamb waves”, *Proceedings of 2008 International Conference on Audio, Language and Image Processing*, pp. 1-4.
- Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P. and Gao, R.X. (2019), “Deep learning and its applications to machine health monitoring”, *Mech. Syst. Signal Process.*, **115**, 213-237. <https://doi.org/10.1016/j.ymsp.2018.05.050>
- Zhou, H.F., Ni, Y.Q. and Ko, J.M. (2011), “Eliminating temperature effect in vibration-based structural damage detection”, **137**(12), 785-796. [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0000273](https://doi.org/10.1061/(ASCE)EM.1943-7889.0000273)
- Zhu, Y., Ni, Y.Q., Jin, H., Inaudi, D. and Laory, I. (2019), “A temperature-driven MPCA method for structural anomaly detection”, *Eng. Struct.*, **190**, 447-458. <https://doi.org/10.1016/j.engstruct.2019.04.004>