Prestressing wire breakage monitoring using sound event detection

Received: 22 March 2023 | Revised: 26 July 2023

DOI: 10.1111/mice.13079

RESEARCH ARTICLE

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

Prestressing wire breakage monitoring using sound event detection

Sasan Farhadi | **Mauro Corrado** | **Oscar Borla** | **Giulio Ventura**

Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Turin, Italy

Correspondence
Giulio Ventura, Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Turin, Italy.
Email: giulio.ventura@polito.it

Funding information
Politecnico di Torino diffuse research grant; Ministry of University and Research (MUR), Grant/Award Number: 20173C478NXFAST-SIMS

Abstract

Detecting prestressed wire breakage in concrete bridges is essential for ensuring safety and longevity and preventing catastrophic failures. This study proposes a novel approach for wire breakage detection using Mel-frequency cepstral coefficients (MFCCs) and back-propagation neural network (BPNN). Experimental data from two bridges in Italy were acquired to train and test the models. To overcome the limited availability of real-world training data, data augmentation techniques were employed to increase the data set size, enhancing the capability of the models and preventing over-fitting problems. The proposed method uses MFCCs to extract features from acoustic emission signals produced by wire breakage, which are then classified by the BPNN. The results show that the proposed method can detect and classify sound events effectively, demonstrating the promising potential of BPNN for real-time monitoring and diagnosis of bridges. The significance of this work lies in its contribution to improving bridge safety and preventing catastrophic failures. The combination of MFCCs and BPNN offers a new approach to wire breakage detection, while the use of real-world data and data augmentation techniques are significant contributions to overcoming the limited availability of training data. The proposed method has the potential to be a generalized and robust model for real-time monitoring of bridges, ultimately leading to safer and longer-lasting infrastructure.

1 | INTRODUCTION

Bridges are essential structures for the smooth operation of transportation networks and significantly affect economic and social development. The number of aged infrastructure networks, particularly bridges, is growing in all developed countries, and the integrity of bridges has become a severe problem. News worldwide reports of bridge collapse cases that, although rare compared to the number of these structures, have deep social and economic consequences. The recent dramatic collapse of the Polcevera bridge in Genova, Italy, in 2018, has evidenced the fragility of some engineering masterpieces and, specifically, the critical effects of prestressing cable degradation. The incident also brought attention to the need for better maintenance and monitoring of these structures to ensure their safety and prevent similar incidents from happening in the future. The corrosion of prestressing cables, which can be promoted by construction defects, exposure to environmental agents, and use of de-icing salts (Bassuoni &

Rahman, 2016; Zhutovsky & Douglas Hooton, 2017), is not easily quantifiable due to the substantial inaccessibility of the cables and the fact that the degradation can be highly localized and, therefore, difficult to spot out. Moreover, it is highly dangerous because it may induce structural collapse without warning signals. Visual inspection is basically unable to assess the status of the prestressing system, and traditional structural health monitoring (SHM) systems show very strong limitations due to the fact that severe loss of prestressing force is required to get measurable effects and, at this point, the safety level can be already compromised. On the other hand, it is practically not possible to carry out detailed investigations on each bridge to detect potential prestressing cable corrosion as these investigations are very expensive and time-consuming, and there are always specific parts that are inaccessible. Therefore, building a system that can efficiently detect the wires failure is an important step toward ensuring long-term operation, safety, and reliability of aging infrastructure networks (Giglioni et al., 2022) as well as preventing costly repairs or replacements.

One approach to detect wire breakage, that is, the acoustic emission (AE) technique, has been extensively studied in material and construction research (Laxman et al., 2023; Yuyama et al., 1995) as it is sensitive to failure and breakage of materials. The release of elastic waves due to the sudden stress variation related to failure can be captured by AE sensors and safely used for event detection on at-risk structures, such as prestressed concrete bridges. However, as pointed out by Yuyama et al. (2007), the application of the AE technique to prestressing wire breakage poses some challenges: signal characteristics produced by wire breakage, wave attenuation during propagation, proper sensor selection and array for sound detection, and wire failure detection under noisy environments. Besides, long-term monitoring produces a large amount of data that require accurate analysis, which can be time-consuming. Therefore, it is necessary to have efficient and reliable data processing and analysis methods to detect potential issues impairing structural safety quickly. One of the most efficient approaches to deal with this issue is utilizing machine learning (ML) and deep learning (DL)-based models. ML and DL are considered subfields of computer science that can build necessary algorithms based on the collected data of certain occurrences (Jordan & Mitchell, 2015). These models can be trained to identify patterns in the data, classify the data into different categories, or make predictions about future events. This can allow for the automated and efficient analysis of large amounts of data, reducing the time and effort required for manual analysis. In recent years, rapid innovations in DL algorithms, improvements in CPU and GPU technology, and the availability of natural or synthetic training data (Gao et al., 2019; Nikolenko, 2021) have provided a considerable impetus for the research and implementation of these techniques in several scientific and engineering fields (Cakir et al., 2015; Farhadi et al., 2022; Luo & Paal, 2023).

Significant progress has been made in the field of damage detection, particularly through the application of AI-based models. Lin et al. (2022) proposed a dynamics-based cross-domain structural damage detection approach using deep transfer learning. Y. Zhang et al. (2022) introduced a novel method for damage detection of nonlinear structures using probability density ratio estimation. Zheng et al. (2022) contributed to this field by proposing a multi-stage semi-supervised active learning framework for crack identification, segmentation, and measurement of bridges. Zou et al. (2022) focused on multicategory damage detection and safety assessment of post-earthquake reinforced concrete structures using DL. Pan and Zhang (2022) introduced a dual attention DL network for automatic steel surface defect segmentation. Furthermore, Amezquita-Sanchez et al. (2018) and Sirca Jr and Adeli (2018) have delved into the critical area of health monitoring of structures. These recent studies collectively demonstrate the active research endeavors in the field of damage detection, providing valuable insights and inspiring further investigations.

The main idea behind the present study is to analyze the ultrasonic signals with the approach known as "Sound Event Detection" (SED), to detect wire breakage on prestressed reinforced concrete bridges. SED has been initially developed from music retrieval and recently has surged attention in various fields using ML and DL approaches, including scene recognition (Chu et al., 2006), surveillance in different environments (Geiger & Helwani, 2015), speech recognition (Mesaros et al., 2021), and audio segmentation (Heittola et al., 2013). Artificial neural networks (ANN), which use multiple hidden layers, have outperformed traditional techniques in the last few years, particularly for SED tasks. Sigtia et al. (2016) compared the classification performance of three different sound-recognition algorithms and highlighted that the ANN model outperforms all other models. To train the ANN algorithm sufficiently, it is essential to extract relevant information; therefore, the AE signals are transformed into a higher level representation. These features, mainly used in acoustic scene classification, speech recognition, and SED, can be extracted from spectrograms (Cakir et al., 2016). Mel-frequency cepstral coefficient (MFCC), as one of the conventional acoustic features, were successfully applied in different learning algorithms (Couvreur et al., 1998; Ito et al., 2009) to provide robust models for sound classification tasks.

However, to the best of our knowledge, this is the first study that applies MFCC to AE signals for detecting wire

breakage in prestressed concrete bridges, an application that poses unique challenges. First, the objective difficulty of collecting real-world data by cutting wires in existing bridges requires enlarging the data set available for training through physics-based data augmentation (DA) approaches, numerical simulations of the breakage process, and/or laboratory tests. Furthermore, there is a need to develop a generalized model capable of detecting wire breakage in bridges that are not part of the training data set. Ideally, this model should exhibit the ability to identify wire breakage in new and unobserved conditions without the necessity for pre-training. Finally, AE signals in the context of wire breakage detection have significant differences with respect to common sound events.

In this research, a novel and tailored approach is presented for wire breakage detection in prestressed concrete beams. The method utilizes a back-propagation neural network (BPNN) combined with input features derived from Mel-frequency cepstral coefficients (MFCCs), specifically adapted and optimized for wire breakage detection. The unique application of these techniques in this context takes into account the physical and spectral characteristics of the wire breakage signals. By carefully adjusting parameters and exploring different configurations, an optimized solution is achieved, enhancing the reliability and performance of the approach. Additionally, DA techniques are employed to address limited training data and improve the model's generalization ability, making it more applicable in real-world scenarios. This research contributes to advancing the field of wire breakage detection by providing a robust and accurate method tailored for prestressed concrete beams.

Compared to other inspection/monitoring techniques, such as radiography (Khedmatgozar Dolati et al., 2023), fiber optics (Hampshire & Adeli, 2000), and so on, the proposed SED model coupled to an AE monitoring system offers unique advantages for early detection of prestressing system degradation. It ensures continuous, automated, and noninvasive monitoring of existing structures, with the ability to detect a single wire breakage in tendons. This cost-effective solution provides essential information for bridge maintenance and safety. The significance of detecting a single wire breakage lies in the fact that corrosion in prestressing wires can be highly localized. Identifying a single wire breakage reveals areas where the entire tendon is undergoing corrosion. The previously introduced inspection and monitoring techniques lack this capability due to limited accessibility (e.g., anchorage zones) or insufficient sensitivity. For example, fiber optics are suitable for new bridges and require substantial breakage to impact local stress in concrete. Additionally, a single wire breakage only affects approximately 0.01% of the first vibration frequency of a typical prestressed bridge beam; therefore, operational modal analysis identification cannot detect such a small change.

The content of the paper is as follows. The experimental process is explained in Section 2. Then, the MFCC technique, different applied DA, and ANNs are proposed in Section 3. Next, metric measurements and model evaluation results are presented in Section 4. Finally, conclusions are drawn in Section 5.

2 | PHYSICAL PROBLEM SETTING

2.1 | Acoustic emission

The AE monitoring technique is based on the acquisition of ultrasonic signals, which typically have frequencies between 20 kHz and 1 MHz. The significance of studying signals in the ultrasonic range lies in that, by doing so, it can effectively isolate the phenomenon to be monitored from the structure operational and ambient background noise. As progressive damage resulting from applied stresses and environmental conditions evolves, local failure may occur. This is characterized by a micro-fracturing process immediately followed by a rapid dislocation that results in the sudden release of energy in the form of transient elastic waves known as AE waves (Grosse & Ohtsu, 2008; Nair & Cai, 2010; Scruby, 1987).

There are three phases for AE: triggering from a weak point, propagating through the volume, and reaching a new equilibrium state. The elastic wave generated by a phenomenon is typically characterized by a steep rising front followed by an exponentially diminishing trend. Eventually, the wave propagates through the material and reaches the surface of the component, where it is detected by piezoelectric sensors. AE events involve a series of waves as a mechanical wave propagation phenomenon. The first part is composed of longitudinal and transverse waves; the second part is the superposition of surface waves and the fraction of directed waves that arrive progressively due to the multiple reflections. AE signals from transducers are then amplified, and pre-processing with the proper filter and threshold settings may eliminate background noise. Finally, signals are recorded and analyzed (RILEM Technical Committee, 2010).

2.2 | Experimental setting

The bridge Alveo Vecchio, located on the A16 Napoli-Canosa highway in Italy, has been selected to collect real-world data. This bridge was set out of service in 2005 following a nearby landslide that precluded its reliability. This condition made it possible to conduct experimental studies under the actual deterioration and constraint

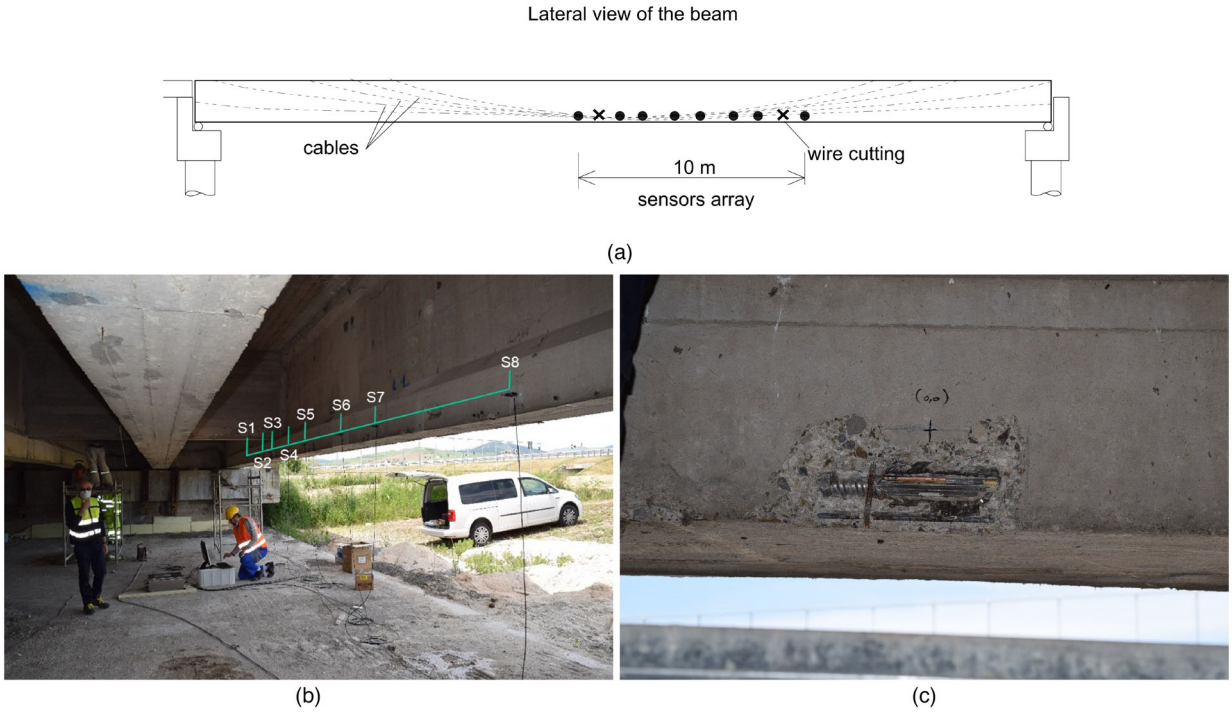

FIGURE 1 (a) Lateral view of the beam; (b) sensors positions on the bridge Alveo Vecchio; (c) close look at the wire cut.

setting to evaluate and assess the bridge safety on-site. Additionally, Alveo Vecchio well represents the typical existing Italian highway bridge. The experimental studies on this bridge, sponsored by Autostrade per l'Italia (ASPI)[1], provide a unique opportunity to assess and clarify the failure and degradation mechanism of prestressed concrete components under operation, thanks to the application of nondestructive and destructive tests and examination on the main structural components.

The Alveo Vecchio viaduct has the following characteristics: two parallel reinforced concrete decks (one for each roadway) and three bays with simply supported I-girders. The deck of each carriageway consists of four longitudinal prestressed reinforced concrete I-girders with a cast-in-place deck slab 20 cm thick; the longitudinal girders are connected by three intermediate diaphragms and two end diaphragms 25 cm thick. The prestressing system consists of 14 post-tensioned cables for each girder. Each cable has 12 parallel wires with a diameter of 7 mm and is placed in a grouted corrugated metal sheath.

In view of the present research, the first preliminary experimental campaign was operated to acquire AE signals from the breakage of post-tensioned prestressing wires in reinforced concrete girders. Failure was triggered by cutting the cross-section of the wire with an electric trimmer until generating a spontaneous tensile breaking. The interdepartmental center SISCON of Politecnico di Torino performed the tests as part of the planned extensive experimental framework that included load tests and many other nondestructive and destructive investigations. The AE signals were acquired using a MISTRAS Sensor Highway III system, considered cutting-edge instrumentation. The system is equipped with eight acquisition channels and specialized software, which allows for the complete acquisition and processing of data. The PK family sensors were selected, which are designed as medium frequency AE sensors, equipped with an integrated, ultra-low noise, low power, and filtered preamplifier that offers a 26 dB amplification. Additionally, it incorporates an integrated Auto Sensor Test (AST) capability, enabling the sensor to both emit and receive signals. This unique feature allows for convenient verification of sensor coupling and performance at any point in time, whether it's before, during, or after a test. The AE signals were captured with an acquisition rate of 2 MHz. Then the signals were time stretched to the audible range for human check of the acquired information, resulting in a fictitious 100 kHz sampling rate (the considered time interval between two samples has been increased, without information loss).

The selected position of the sensors (S1–S8) was on the lower edge of the girder (Figure 1a,b) to be able to evaluate the AE signal transformation and attenuation through the structural components, including concrete. This is essential to accurately acquiring and interpreting the acoustic

[1] the experimental studies are part of a collaboration agreement between the "Ministry of Infrastructure and Sustainable Mobility" (MIMS), ASPI, Università di Trento, Politecnico di Torino, and other Italian Universities.

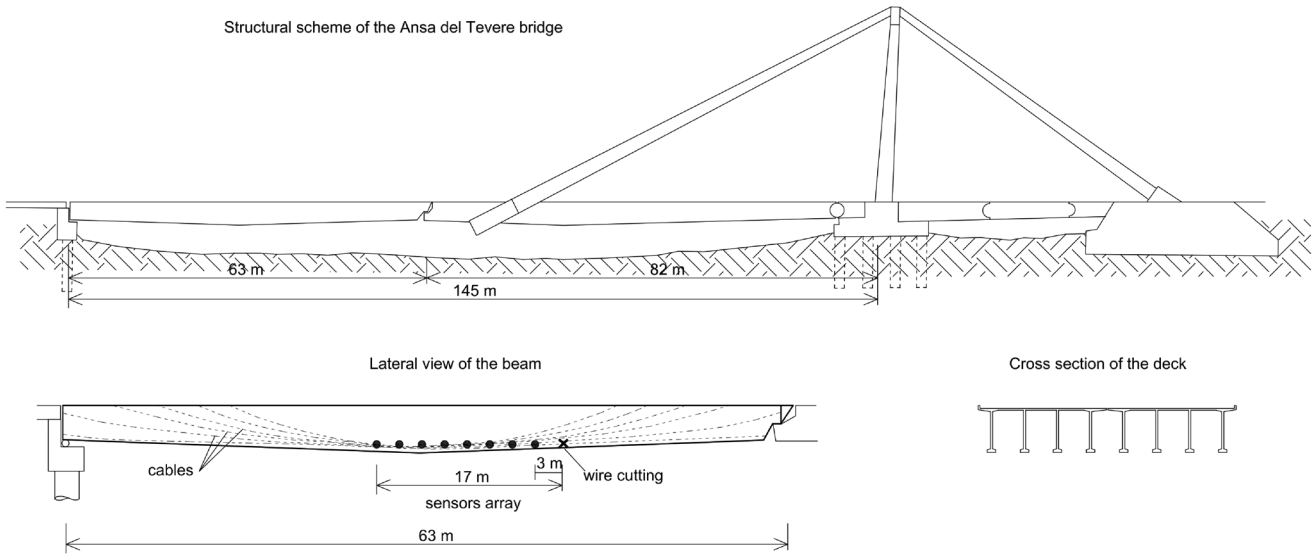

FIGURE 2 Structural scheme of the Ansa del Tevere bridge located in Roma, Italy.

signals and verifying the ability to detect the wire cut despite a substantial distance between the breaking point and the sensor. Eight transducers were used; two of them were broadband (positioned S1 and S8), while the rest were resonant. The cutting operation was performed on two beams (called T3 and T4), with eight cuts per beam and 16 in total. For each specific cut scenario, different thresholds were applied to the sensors involved. For example, when cutting occurred between sensors S1 and S2, thresholds of 75 dB were used for S1 and S2, 60 dB for S3, 55 dB for S4, S5, and S6, and 25 dB for S7 and S8. Remarkably, no AE events were recorded during the cutting operations until the tensile breaking. Concrete core drilling was performed in an adjacent girder, but this was neither detected as an AE event. This has shown a good insensitivity of the AE technique to external disturbances while gathering wire breakage events.

The second phase of the experimental tests took place on the "Ansa del Tevere" viaduct, also known as "Ponte Morandi" or "Viadotto Morandi." This cable-stayed motorway suspension bridge is located in Roma, Italy, on the A91 Roma-Fiumicino highway in the Magliana district. The bridge has a reinforced concrete structure and suspension cables embedded in square concrete castings (see Figure 2). The western cables support the prestressed concrete deck, while the eastern cables are anchored to reinforced concrete caissons to act as a counterweight. The main span consists of a 74-m long cantilever multi-cell box girder and a 63-m long twin deck with eight prestressed reinforced concrete I-girders with a cast-in-place deck slab 16 cm thick and six intermediate diaphragms. The prestressed reinforced concrete I-girders are simply supported on a pile, and the cantilever part is through a Gerber hinge. Although different types of cables (with straight and twisted wires) have been cut in different structural elements, only the results from a parallel wire tendon placed in one of the I-girders are included in this study. This choice was made for better consistency with the data obtained from Alveo Vecchio. In particular, the prestressing system of the I-beams consists of 18 cables having 21 parallel wires each. The diameter of a single wire is 7 mm. The arrangement of the eight AE sensors along the beam is shown in Figure 2.

3 | METHODS

In this section, a conceptual description of MFCC feature extraction is given. Then, the DA techniques are explained briefly and, finally, proposed ANN model is discussed.

3.1 | Feature extraction

MFCCs are one of the most dominant and prevalent features used for SED, and one of the reasons for their success is the ability to represent the signal spectrum compactly. Figure 3 indicates the required steps to extract MFCCs from the original signals. This process can be explained in detail as follows (Logan, 2000):

1. **Frame blocking and windowing of the signal**

   The AE signals are considered nonstationary signals; consequently, to have stable acoustic features, it is essential to consider them over a sufficiently short period. Therefore, the original waveforms are divided

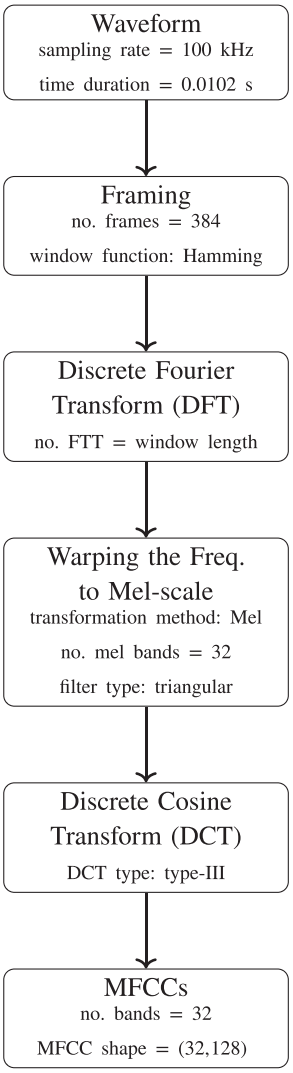

FIGURE 3 Process to extract MFCC features.

into smaller frames at fixed intervals using windowing functions (Blackman & Tukey, 1958) such as Hamming, Hann, and Bartlett. In general, the purpose of applying window functions is to smooth the frames, eliminate the edge effects, and enhance the harmonics in the signals while using the DFT on the signals (Logan, 2000). In addition, the window function secures that the signal ends close to zero. The following formula represents Hamming window function (Rao & Manjunath, 2017) which was used in this study:

$$w[k] = 0.54 - 0.46\cos\left(\frac{2\pi k}{K}\right) \quad k = 0, \ldots, K-1 \tag{1}$$

where K is the number of data points in each frame.

2. **Discrete Fourier Transform**

At the end of the initial phase, a vector of cepstral features is generated for each created frame. In this step, the Discrete Fourier Transform (DFT) of each frame

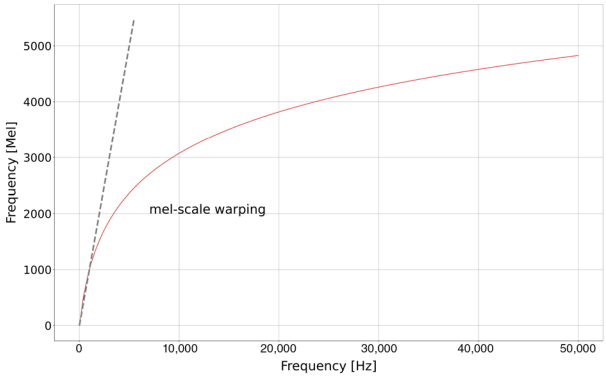

FIGURE 4 Scaled Mel frequency versus physical frequency.

is computed using Equation (2), and the log of the amplitude spectrum is obtained as the perceived signal loudness has a logarithmic behavior (Logan, 2000).

$$X[k] = \sum_{n=0}^{N-1} x(n)e^{\frac{-j2\pi nk}{N}} \qquad 0 \le k \le N-1 \tag{2}$$

where N is the number of points used to compute the DFT.

3. **Warping the frequencies to the Mel-scale**

In this step, the Mel spectrum is computed by passing the computed frequency content Equation (2) through a set of band-pass filters known as the Mel-filter bank. This is essential for capturing the spectral characteristics of the signal relevant to wire breakage detection. Mel is an abbreviation of melody and a measurement unit based on human auditory perception (Stevens & Volkmann, 1940; Stevens et al., 1937). As indicated in Figure 4, Mel-scale has a linear behavior below 1000 Hz, whereas, beyond this range, it increases logarithmically and is highly compressive (Rao & Manjunath, 2017). A good approximation of the Mel-scale can be derived from the physical frequency as follows (Beigi, 2011):

$$f_{\text{mel}} = \frac{1000}{\log(2)}\log\left(1 + \frac{f}{1000}\right) \tag{3}$$

where f_{mel} represents Mel-frequency and f denotes physical frequency. The filter banks apply to both the frequency and time domains; however, for MFCC calculation, the filters are typically implemented in the frequency domain. The most typical filter is triangle-shaped. Mel-spectrum of the magnitude spectrum $X[k]$ can be obtained by multiplying $X[k]$ by each triangular Mel weighting filter as follows:

$$S[m] = \sum_{k=0}^{N-1}[|\,X[k]\,|^2\,H_m[k]] \qquad 0 \le m \le M-1 \tag{4}$$

where M is the total number of triangular filters usually selected into 20–40 (Logan, 2000). The reason that conventionally 20–40 filters are used is that it provides a good trade-off between the resolution and complexity. If too few filters are selected, some important information (features) can be lost or smoothed out; in contrast, if too many filters are used, it will be computationally expensive, and some features can be correlated. It is worth noting that the number of filters can also depend on other factors, such as the sampling rate and frequency range of the signals (Pan & Zhang, 2022). $H_m[k]$ is the weight given to the kth energy spectrum bin contributing to the mth output band and expressed as follows:

$$H_m[k] = \begin{cases} 0 & k < f(m-1) \\ \dfrac{2(k - f(m-1))}{f(m) - f(m-1)} & f(m-1) \le k \le f(m) \\ \dfrac{2(f(m+1) - k)}{f(m+1) - f(m)} & f(m) \le k \le f(m+1) \\ 0 & k > f(m+1) \end{cases} \tag{5}$$

where $f(m)$ can be expressed as:

$$f(m) = \left(\frac{N}{f}\right) f_{\text{mel}}^{-1}\left(f_{\text{mel}}(f_l) + m\frac{f_{\text{mel}}(f_h) - f_{\text{mel}}(f_l)}{M+1}\right) \tag{6}$$

where f_l and f_h indicate the lowest and highest frequencies, respectively.

4. **Extract MFCCs using inverse Discrete Cosine Transform**

The inverse DCT or DCT-III reported in Equation (7) is applied to the transformed Mel frequency coefficients and computes cepstral coefficients.

$$C[n] = \sum_{m=0}^{M-1} \log(S[m]) \cos\left(\frac{\pi n(m - 0.5)}{M}\right)$$
$$n = 0, 1, 2, \ldots, C-1 \tag{7}$$

C is the number of Mel bands, and $C[n]$ are the cepstral coefficients. Evaluating the inverse DCT from the signal $X[k]$ gives the representation of the energy content of the signal. One advantage of MFCCs is their robustness toward noise and spectral estimation errors under various conditions (Balsamo et al., 2014). The algorithm used to extract the coefficients is summarized in Algorithm 1

ALGORITHM 1 A pseudo-code algorithm for MFCC extraction

Input: signal
Output: mfcc_features - The extracted Mel-frequency cepstral coefficients
Define the *frame_length* as the desired duration of each frame in seconds
Define the *hop_length* as the desired time shift between consecutive frames in seconds
frames ← *SplitInFrames(signal, frame_len, hop_len)*
mfcc_features ← []
for each *frame* in *frames* **do**
  spectrum ← *FFT(frame)*
  power_spectrum ← *PowerSpectrum(spectrum)*
  mel_spectrum ← *MelFilterbanks(power_spectrum)*
  cepstral_coefficients ← *DCT(mel_spectrum)*
  mfcc_features.append(cepstral_coefficients)
end for
return mfcc_features

3.2 | Data augmentation

The number of parameters, such as weights and bias, to train an ANN model is exceptionally high; therefore, to build a generalized model, a large number of training data sets is required for each class to cover acoustic class variability based on the model complexity (Takahashi et al., 2016; C. Zhang et al., 2021). Moreover, many possible sound combinations are limited or cannot be present in the recorded data, making generalization more difficult. There are multiple ways to maximize the model performance on the test data set; however, DA plays a key role and is a prevalent way to increase training sets for learning algorithms artificially (H. Zhang et al., 2017). In theory, an accurate DA allows to the elimination of the performance differences between the train and test sets (Chun et al., 2022).

DA techniques used for SED can be varied from time stretching and dynamic range compression for essential signal augmentation to mixup and block mixing for complex ones (Mesaros et al., 2021). In the present study, polarity inversion, mixup, and time-shifting were used to enhance the model performance.

3.2.1 | Polarity inversion

This is a simple DA method that flips the waveform by multiplying it by −1. The assumption is that a signal's polarity does not affect its acoustic content. Therefore,

flipping the waveform will not change the information content of the signal but create a new sample that can be added to the training data set. Polarity inversion can be useful for increasing the variability of the data set and making the model more robust to changes in signal polarity. In this study, the polarity of a randomly selected subset of the training data was inverted to create additional training data.

3.2.2 | Mixup

Mixup DA introduced by H. Zhang et al. (2017) was employed to generate new training samples by linearly interpolating existing ones. An augmented signal s_{mix} can be generated as follows:

$$s_{\text{mix}} = \lambda x_i + (1 - \lambda) x_j \tag{8}$$

where, x_i and x_j are the original source signals from the same class, and λ is a value between 0 and 1 randomly extracted from a uniform or beta distribution. This formula results in a new mixed sample that has some characteristics of both original signals, introducing more variation in the training data.

3.2.3 | Time-shifting

To further enhance the diversity of the environmental noise training data set, time-shifting was employed. This method involves shifting the signals by a defined time interval to generate new samples. Each signal has been divided into eight segments and shifted each segment by a random number of samples between 0 and the length of the signal. This technique produced eight new samples for each original signal. By applying time-shifting, the model can become more resilient to slight variations in the time alignment of the ultrasonic signals, which can occur due to various environmental factors. This technique can help improve the overall performance and accuracy of the model by introducing additional variation in the training data.

All these techniques can increase the data set diversity and make the model more robust to sound event variations. However, it is worth noting that when applying DA techniques, it is essential to ensure that the modifications are realistic and do not introduce new classes or nonphysical sound events.

3.3 | Back-propagation neural network

This section briefly explains the BPNN utilized for binary classification. BPNN is a supervised learning algorithm that consists of input, hidden, and output layers, as illustrated in Figure 5. In this study, the input layer receives the MFCC values, while the output layer represents the source signal class, namely wire breakage or environmental noise. The network is trained to predict a binary output, with a value of 1 indicating wire breakage and a value of 0 indicating environmental noise. The hidden layers receive the weighted combinations of these variables and employ activation functions to give the output targets. The learning algorithm is based on a gradient-based propagation technique, with an error function used to compute the error between the model output and observed data to update the weight values in each iteration (Wang et al., 2022). However, MacKay (2019) proposed employing the "cross-entropy" error function instead of the sum-of-squares for a classification problem leads to faster training and improved generalization. The error function is the negative log-likelihood of the true labels given the predicted probabilities, computed as in Equation (9). The function output is a scalar value reflecting the model loss or error. The weights and biases will be adjusted through the back-propagation steps to decrease the loss value and improve the model performance (Chollet, 2018). The formula for binary cross-entropy loss is given by Murphy (2013):

$$CE = -\frac{1}{S}\sum_{i=1}^{S} y_i \log p(y_i) + (1 - y_i) \log (1 - p(y_i)) \tag{9}$$

where S represents the total number of samples in the data set. Each sample i is associated with a true label y_i, where y_i takes the value of 0 for noise and 1 for breakage. The predicted probability of the positive class (breakage) is denoted as $p(y_i)$. Additional descriptions of ANN models can be found in the literature (Bishop, 2006; Goodfellow et al., 2016).

4 | EXPERIMENTS AND ANALYSIS

This section explains data collection and implementation, model hyperparameters, metrics measurements used for the model evaluation, classification results and overall evaluation.

4.1 | Data set and implementation

As this is the first study to classify wire failure as sound event for SHM using AE signals, a rigorous data selection procedure is required to ensure that the data set is representative of the important events that may occur during the bridge lifetime. Besides, there is currently no available

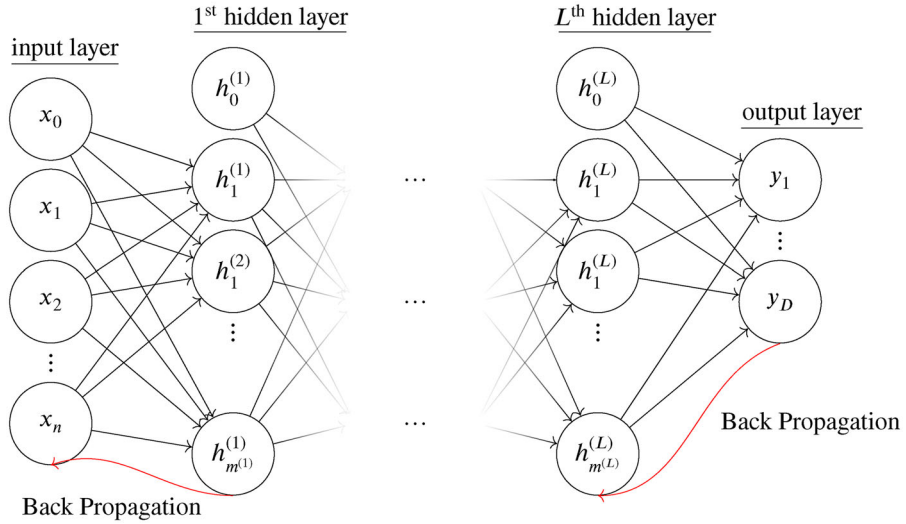

FIGURE 5 Structure of back-propagation neural network.

data set of wire breakage events; therefore, generating a data set suitable for training the event detection models was a key element. The primary data set collected from the Alveo Vecchio bridge contains 244 acoustic signals, including 128 wire cut signals and 116 environmental noise signals. The wire cut signals were accurately recorded by fixing an optimal threshold value on amplitudes to avoid recording any noise during the cutting operation. This is important to ensure that the data set only contains relevant information and that the data are reliable. Strong label signals make the available data set more robust, and augmenting different classes would be easier and more efficient. It will also minimize the possibility of errors during the detection of sound events in real-time monitoring. Multiple channels help capture sound events from different positions, improving detection and classification accuracy. As previously evidenced, the pre-processing phase involved stretching in time the signals to adjust their frequency range from ultrasonic to audible frequencies and, thus, consistently applying the Mel-spectrum representation, originally developed to highlight the human perception of a sound. The re-scaling process resulted in a final sampling rate of 100 kHz with a frequency range of 0–50 kHz.

In certain scenarios, it can be challenging to distinguish between wire breakage signals and environmental noise based on their characteristic parameters, such as amplitude and energy alone. Figure 6a,b depicts the time-domain representations of wire breakage signals, while Figure 6c represents the environmental noise signal. In the signal in Figure 6b, the absolute energy, amplitude, and risetime values are significantly higher with the values of 3.8×10^7 aJ, 1.5 V, and 704 μs, respectively, compared to the environmental noise signal in Figure 6c, which

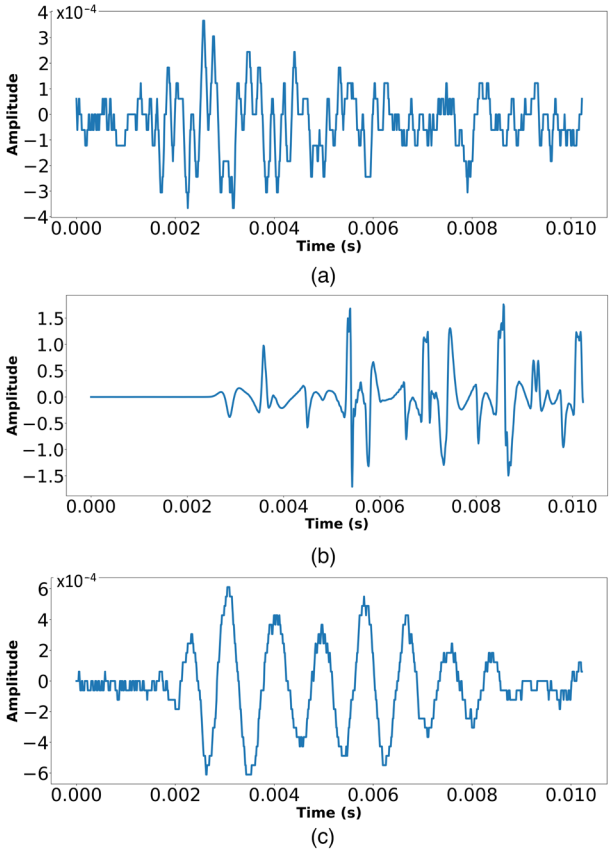

FIGURE 6 Time-domain representation: (a, b) wire breakage signal; (c) environmental noise signal.

has values of 6.2 aJ, 6×10^{-4} V and 49 μs. Conversely, in Figure 6a, representing a wire breakage signal detected by a sensor far from the cut, the corresponding energy, amplitude, and risetime values are lower with the values of 0.54 aJ, 3.8×10^{-4} V and 47 μs, respectively, compared

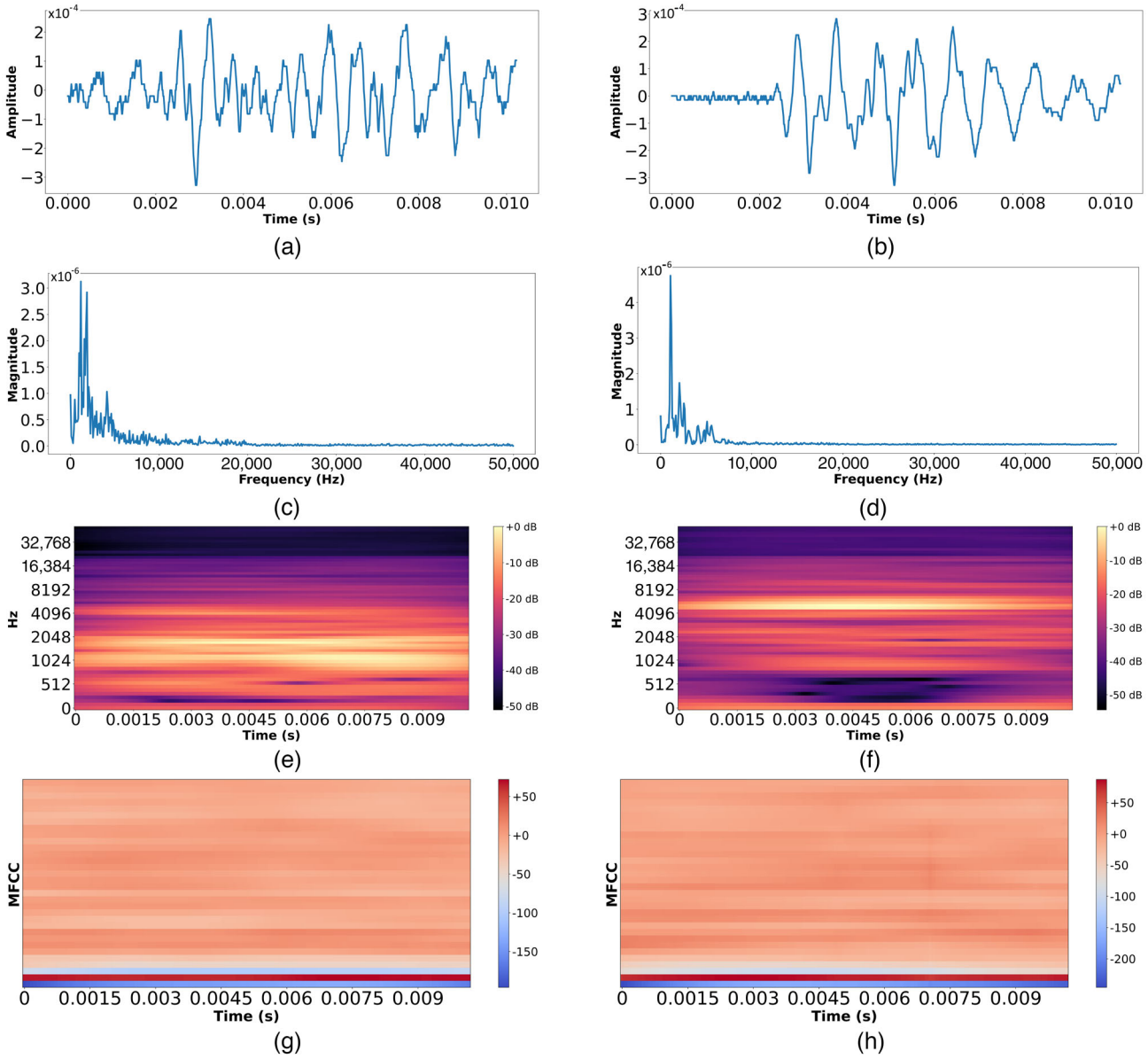

FIGURE 7 (a, b) Time-domain waveform visualization; (c, d) frequency-domain analysis; (e, f) Mel-spectrum representation; (g, h) MFCC characterization for wire breakage and environmental noise, respectively.

to the wire breakage signal in Figure 6c. These observations highlight the challenge of discerning differences between wire breakage signals and environmental noise based solely on these parameters. On the other hand, the signals represented in Figure 6a and c may contain distinct patterns and frequency components that can be effectively captured and analyzed through advanced signal processing techniques such as Mel-spectrogram and MFCC.

Figure 7 provides different representations of the analysis conducted on wire breakage and environmental noise signals. The time-domain waveform representations are depicted in Figure 7a,b, indicating the temporal characteristics of the preprocessed breakage and noise signals. Moving to the frequency domain, Figure 7c,d presents the frequency-domain analysis, revealing the spectral content of the signals. In order to extract meaningful features for further analysis, the sound samples were converted into MFCCs using Python and the SciPy library, as described in Section 3. The choice of FFT length is critical, and it was ensured that it is at least the same length as the window to provide a high-resolution representation and sufficient information. For this study, 32 filter banks with 256-FFT points were utilized, resulting in compact features comprising 384 frames. An example of the Mel-spectrum and MFCC representation is presented in Figure 7e,f. It is important to note that the selection of the FFT length was based on the signal sampling rate and the desired

TABLE 1 Distribution of original and augmented data set for training and testing: Alveo Vecchio bridge.

Data type	Number of samples	Percentage
Original (total)	244	–
Original (wire breakage)	128	52.50%
Original (environmental noise)	116	47.50%
Original (training set)	195	80.00%
Original (test set)	49	20.00%
Augmented (total)	2706	–
Augmented (wire breakage)	1374	55.00%
Augmented (environmental noise)	1332	45.00%

TABLE 2 Optimized hyperparameters: Key factors for trained model performance on both original and augmented data set.

Hyperparameter	Selected parameters
Input size	32×128
Number of hidden layers	5
Number of epochs	500
Activation function	leaky-relu
Learning rate	9.5E-7
Optimizers	Nadam
Initializer	Glorot
Number of neurons (1st layer)	1000
Number of neurons (hidden layers)	90

frequency resolution, ensuring an appropriate representation of the underlying signal characteristics. These representations serve as valuable inputs for subsequent analysis and classification tasks, enabling the effective characterization and differentiation of wire breakage and environmental noise signals.

To ensure the validity of the proposed approach and to build a baseline model, the model was initially trained solely on the data collected from the Alveo Vecchio bridge, which consist of 244 signals, including 128 wire breakage events and 116 environmental noise events. The train-test set has been split into 80%–20%. Then, to increase the sample size and improve the model performance, a DA approach was employed to both train and test sets, keeping the proportion 80%–20%, which resulted in 2706 total events, including 1374 wire cuts and 1332 environmental noise events. The DA approach included time-shifting for environmental noise and polarity inversion and mixup techniques for both wire breakage and environmental noise events. Table 1 presents the distribution of the original and augmented data sets for training and testing on the Alveo Vecchio bridge. Moreover, due to the uncertainty of materials and structures, evaluating the models on different structures with specific geometry and characteristics is vital. To address this concern, the secondary data set collected from the Ansa del Tevere bridge was used to evaluate the trained model capability for wire breakage detection on a different structure. The data collected from Ansa del Tevere were solely used as an unseen test data set.

4.2 | Model parameters

To optimize the training process of the BPNN models, we utilized MFCCs with a consistent size of 32×128. This standardized format ensures efficient data processing capabilities of the models (Gao et al., 2019). For each model, we carefully selected specific hyperparameters, which are listed in Table 2. Given the potentially slow and computationally demanding nature of training an ANN model with numerous parameters, various strategies were adopted to enhance the model performance, as described in the following.

- **Feature extraction**: To solve this problem, MFCCs were extracted to train the model, which decreased the size of training from $[x * x]$ traditional spectrum to $[k * n]$, where $k, n < x$.
- **Data set size**: DA techniques were implemented to increase the size of the data set.
- **Initialization**: Glorot initialization (Glorot & Bengio, 2010) was set to avoid vanishing and exploding gradient problems.
- **Regularization**: To avoid overfitting and enhance the model performance, some regularization techniques, including batch normalization (BN) proposed by Ioffe and Szegedy (2015), and drop-out initially introduced by Hinton et al. (2012) were applied.
- **Optimization**: Nadam optimizer (Dozat, 2016) was applied to boost solutions convergence.
- **Hyperparameters**: Randomized search CV approach was utilized to find the optimal hyperparameters (Table 2).
- **Early stopping**: Callback was used to interrupt the training processes when the selected validation metrics stopped improvements.

Moreover, to avoid the risk of sampling bias and information leaks (Chollet, 2018), cross-validation using a stratified shuffle technique has been employed where n-split set to 10 and the validation size to 0.1. This technique merges shuffle split and stratified fold split and allows for a more robust evaluation of the model. The model

evaluation metric has been set to "loss" to assess the model performance.

4.3 | Evaluation criteria

Evaluating the performance of trained DL models is an essential step in assessing their effectiveness. Therefore, computational metrics were employed to asses proposed models by comparing the predicted output and ground-truth labels. The comparison between these two values was obtained using true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Accuracy (Ac), Precision (Pr), Recall (Re) or Sensitivity (Se), Specificity (Sp), and F1-score (F1) were employed as common methods in binary classification.

In addition, Matthews Correlation Coefficient (MCC) introduced by Matthews (1975) was implemented using Equation (10), which computes the coefficients for binary classes. The values for MCC range in [−1, +1], with the best scores +1. It is the only binary classification metric that considers all the confusion matrix parameters and generates a high score when the model can predict positive and negative classes.

$$\mathrm{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \tag{10}$$

When there are no positive or negative measurements, MCC will be undefined. The average conditional probability (ACP) (Baldi et al., 2000) was used to face this problem and capture both specificity and sensitivity:

$$\mathrm{ACP} = \frac{1}{4}\left[\frac{TP}{TP + FN} + \frac{TP}{TP + FP} + \frac{TN}{TN + FP} + \frac{TN}{TN + FN}\right] \tag{11}$$

For the model to be reliable and applicable to real-world scenarios, besides the accuracy, its performance in minimizing both FP and FN is of paramount importance. Obviously, a high number of FN (breakage signals identified as noise) has a negative impact on the model reliability. However, the number of prestressing wires in a concrete beam is always very large; accordingly, missing a single wire breakage is not a critical issue. Therefore, an acceptable probability of having FNs could range from 5% to 10%. Structural collapse requires the breakage of multiple wires, making the joint probability of missing all events extremely low. In a real application scenario, the detection of a single wire breakage triggers human intervention for a further and more invasive inspection to better assess the condition of the structure. Common interventions require removing a portion of concrete in the region where the damage has been localized and performing a visual inspection of the prestressing tendon. However, such an operating protocol makes the detection of FP (noise signals identified as breakage) more critical. As the noise signals recorded by the AE system in a real case are much more than the breakage signals, even a very low probability of having FP (e.g., 1%) may trigger the human intervention every few days or weeks, which is certainly not acceptable. Therefore, although high accuracy is important, a primary focus is on minimizing FP, thereby achieving a high Recall/Sensitivity score.

4.4 | Classification result

This section provides an overview of the results obtained from our proposed BPNN models for wire breakage detection, utilizing both the original and augmented data set. Figure 8 presents an overview of the workflow in this approach. BPNN algorithm with two different regularization techniques, namely, BN and dropout, was implemented using the Keras library—a high-level neural network API—on the TensorFlow platform. The classification performance of these models is discussed in detail in this section.

To assess the trained models performance, initially, they were evaluated on the test set from the Alveo Vecchio bridge, which was used during the training phase. This enabled evaluation of the model performance on the data it was trained on and overall reliability assessment. Furthermore, the trained models were evaluated on an unseen data set from the Ansa del Tevere bridge to assess their ability to generalize on new scenarios and real-world applications.

To gain a deeper understanding of the performance of the proposed models, confusion matrices were employed to evaluate their effectiveness. The sequential model achieved an accuracy of 93.85% on the Alveo Vecchio test set indicating its ability to make correct wire breakage predictions. Moreover, it could attain an F1-score of 93.00% and MCC score of 0.88, indicating its efficiency in striking a balance between precision and recall (Figure 9a). Similarly, this model with BN and dropout regularization could achieve an overall accuracy of 98.00%. This enhancement was accompanied by an F1-score of 98.30% and an MCC score of 0.96, illustrating the efficacy of the proposed regularization techniques in improving model performance

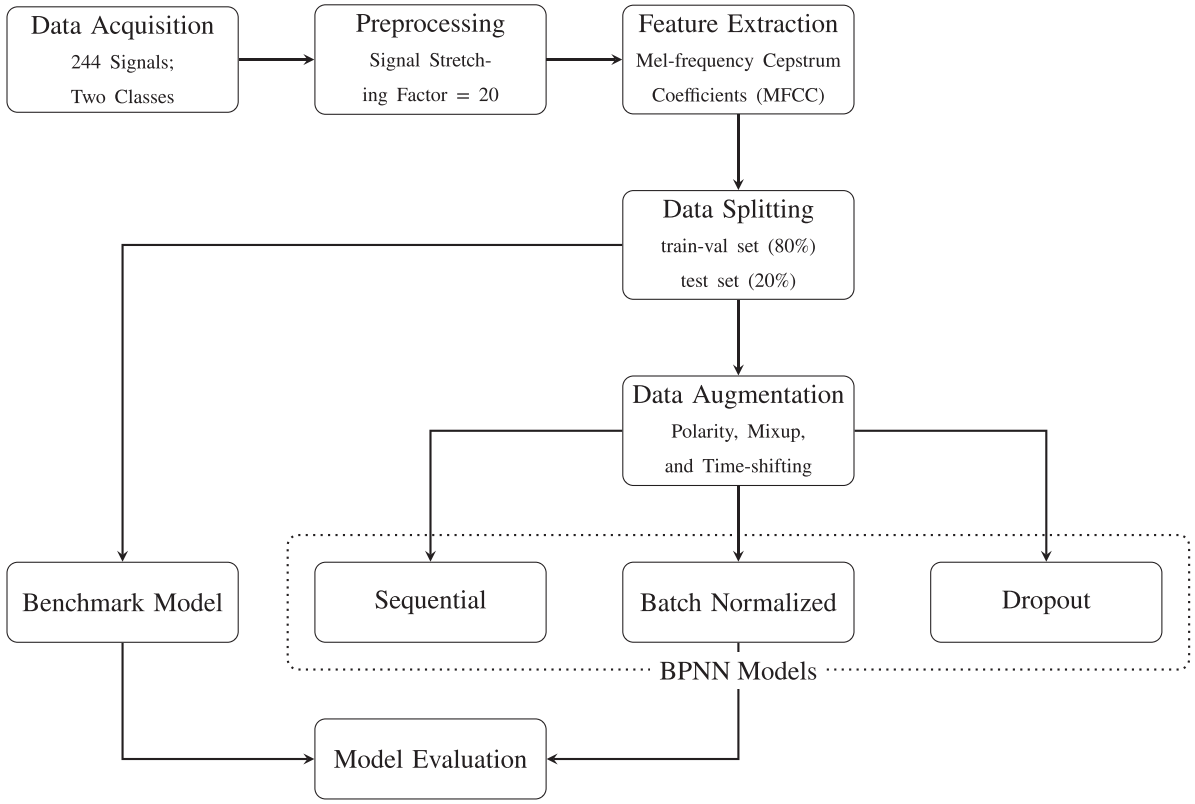

FIGURE 8 The strategy workflow for sound event detection approach to identify prestressing wire breakage in bridges.

TABLE 3 Performance metrics comparison for event detection on test set.

Models		Accuracy	Precision	Recall	Specificity	F1-score	MCC	ACP
Alveo Vecchio	Baseline	95.50	100.00	91.20	100.00	95.40	0.91	0.95
	Sequential	93.85	100.00	87.00	93.00	93.00	0.88	0.94
	Batch	98.00	100.00	96.50	100.00	98.30	0.96	0.98
	Dropout	98.00	100.00	96.50	100.00	98.30	0.96	0.98
Ansa del Tevere	Baseline	67.00	33.30	30.80	79.95	32.00	0.09	0.55
	Sequential	61.00	23.10	23.10	73.70	23.00	0.03	0.48
	Batch	78.50	75.00	23.00	97.40	35.30	0.35	0.68
	Dropout	82.35	70.00	54.00	92.10	61.00	0.58	0.75

(Figure 9b and c). In the experimentation process, all models were trained for 250 epochs to thoroughly assess their capabilities and convergence patterns (Figure 10). It is worth noting that early stopping was employed as a precautionary measure to prevent overfitting. A patience value of 5 and monitoring the loss function were utilized as criteria for early stopping.

The final results presented in Table 3 indicate the impact of utilizing augmented data on model performance. It is noteworthy that the average performance metrics achieved by the augmented data models surpass those of the sequential model without augmentation. Among the models employing regularization techniques, the sequential model with regularization emerged as the most effective, showcasing the highest scores across all metrics. The F1-scores of 98.30% attained by these models illustrate their ability to strike a balance between precision and recall, encompassing both positive and negative classes accurately. Moreover, these models exhibited the highest MCC and ACP scores, recording values of 0.96 and 0.98, respectively. These scores reflect the model outstanding predictive power, especially in identifying the highest positive and negative classes. These findings highlight the fundamental value of augmented data and the efficacy of employing regularization techniques in enhancing the performance of wire breakage detection models.

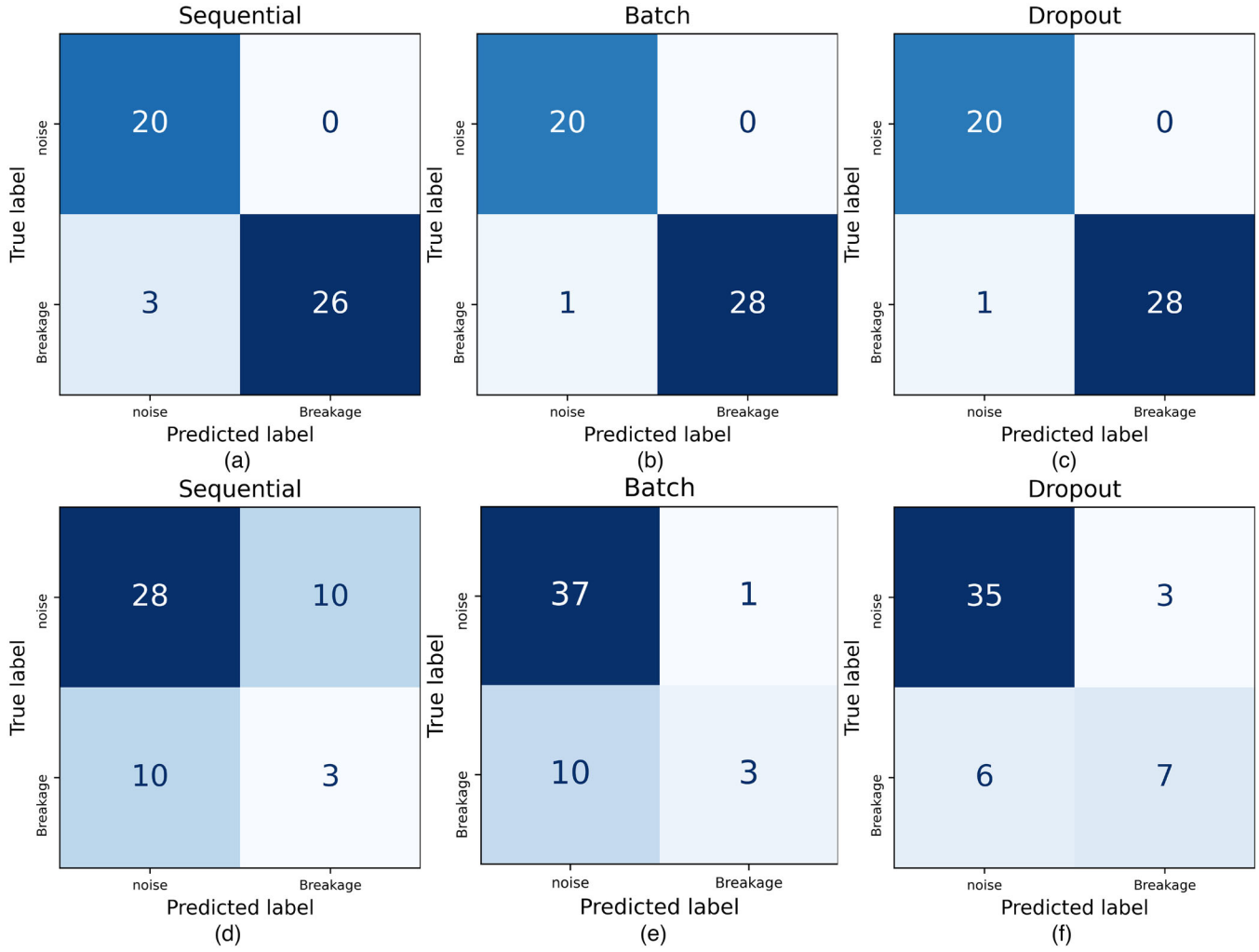

FIGURE 9 Confusion matrices for testing results on the Alveo Vecchio bridge (a–c) and Ansa del Tevere bridge (d–f).

In the second step of our analysis, the trained models were applied to detect wire breakage signals in the Ansa del Tevere bridge. The number of wire cut signals collected in this bridge was 13, while there were 38 environmental noise signals. The evaluation of the models, as depicted in Table 3, revealed a decrease in their performance across all measured metrics when compared to the Alveo Vecchio bridge test set. A robust model should have high sensitivity and specificity to minimize false detections in breakage and noise signals. Nevertheless, it is crucial to acknowledge that the decline in model performance may stem from various underlying factors, extending beyond the only consideration of their robustness. For instance, a practical issue that was evidenced in the Ansa del Tevere bridge and that makes it different from the Alveo Vecchio bridge is poor grouting of the sheath. This issue can affect how the ultrasonic wave propagates through the structure and, therefore, its shape and amplitude when detected by the sensors. The confusion matrices presented in Figure 9d–f provide a better understanding of the model performance. Notably, the sequential model, devoid of any regularizers, just identified three wire cut signals. On the other hand, the model with BN obtained the highest TN values, correctly detecting 37 out of 38 environmental noise signals, and achieved the highest specificity score of 97.40%.

On the other hand, the dropout model demonstrated superior performance compared to the other models by successfully detecting 7 out of 13 wire breakage signals, achieving accuracy and F1-score of 82.35% and 61.00%, respectively. The MCC values for the sequential models, both without and with augmentation, were close to zero, indicating random predictions and a lack of agreement between the true labels and the predicted labels. These findings suggest that the dropout model holds the most promise for effectively detecting wire breakage signals in the Ansa del Tevere bridge data set. However, further fine-tuning may be necessary to reduce the number of false positives, thereby enhancing the model's precision in accurately identifying genuine wire breakage instances. These results underscore the significance of employing

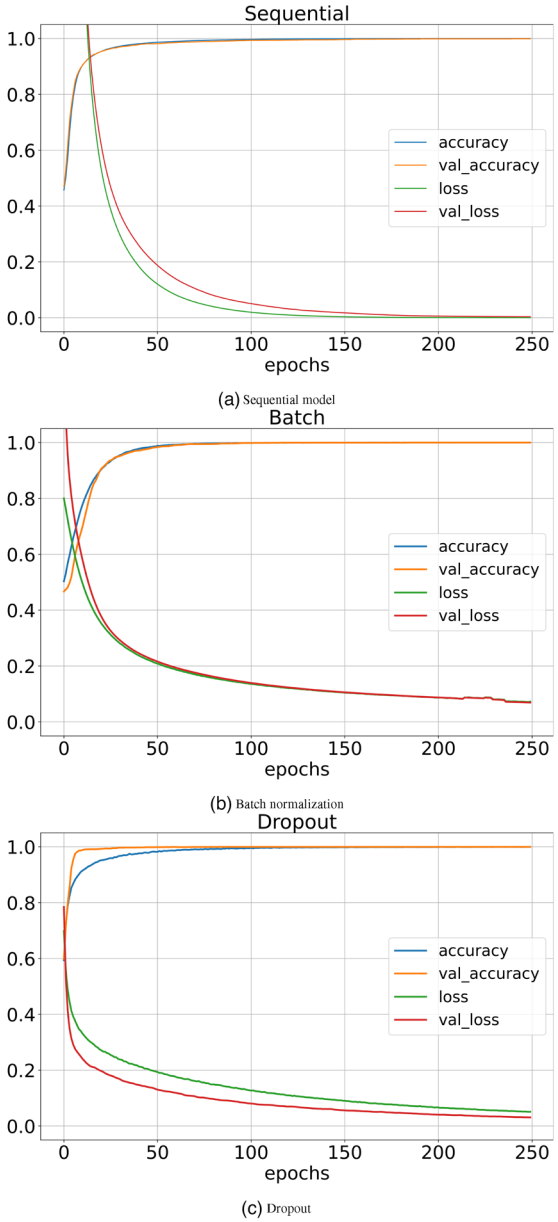

FIGURE 10 Learning curves using augmented data set.

regularization techniques in this task to improve overall model performance and reliability.

5 | CONCLUSION

This study proposes a novel approach for detecting and classifying wire breakage and environmental noise signals in prestressed concrete bridges using BPNN models. The proposed models utilized MFCC extracted from AE signals as input data, and different DA techniques were employed to address the scarcity of real data for training ANN algorithms. Polarity inversion, mixup, and time-shifting effectively augment wire cut and environmental noise signal data sets. These techniques preserve the physics of the signals, closely resembling the original data and improving performance in classification and detection. The successful application of MFCC to detect wire breakage is particularly noteworthy. It reduces data dimensionality while effectively representing spectral features in a compact form, which is critical for model training and generalization. Two regularizers were applied to improve the BPNN algorithm generalization performance, with the dropout model proving to be the most promising one.

For the future expansion of this research, some more advanced ML algorithms are planned to be explored, such as Neural Dynamic Classification Algorithm (Rafiei & Adeli, 2017), Ensemble Learning (Alam et al., 2020; Laxman et al., 2023), and Self-Supervised Learning (Rafiei et al., 2022). These models have shown promising results in different applications and may improve the reliability of wire breakage detection.

Despite the promising outcomes of the current study, improving the model reliability for different bridges is needed. Ultrasonic waves are varied by structure and materials, making signal characteristics somehow unique to each bridge, posing a challenge for having a universal model for wire cut detection. To address this issue, future works can be directed toward developing a more diverse data set that includes a broader range of bridge types. Although this study was focused on wire breakage, the proposed SED approach can be extended to identify a wider range of structural damage mechanisms, becoming an essential and effective tool for the permanent safety monitoring of new or aging infrastructures.

ACKNOWLEDGMENTS

Research Grant "XFAST-SIMS: Extra fast and accurate simulation of complex structural systems" by the Italian Ministry of Education, University and Research and research support by Politecnico di Torino with the diffuse research grant initiative are gratefully acknowledged. The authors are grateful to ANAS S.p.A. and Ing. Alberto Gennari Santori of Cesano Research Center for their availability and support in collecting Ansa del Tevere bridge data. The authors are also grateful to Autostrade per l'Italia S.p.A. and the University of Trento for their availability and support in collecting Alveo Vecchio bridge data.

REFERENCES

Alam, K. M. R., Siddique, N., & Adeli, H. (2020). A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications*, *32*(12), 8675–8690.

Amezquita-Sanchez, J., Valtierra-Rodriguez, M., & Adeli, H. (2018). Wireless smart sensors for monitoring the health condition of civil infrastructure. *Scientia Iranica*, *25*(6), 2913–2925.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics*, *16*(5), 412–424.

Balsamo, L., Betti, R., & Beigi, H. (2014). A structural health monitoring strategy using cepstral features. *Journal of Sound and Vibration*, *333*(19), 4526–4542.

Bassuoni, M., & Rahman, M. (2016). Response of concrete to accelerated physical salt attack exposure. *Cement and Concrete Research*, *79*, 395–408.

Beigi, H. (2011). *Fundamentals of speaker recognition*. Springer US.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Blackman, R. B., & Tukey, J. W. (1958). The measurement of power spectra from the point of view of communications engineering: Part I. *The Bell System Technical Journal*, *37*(1), 185–282.

Cakir, E., Heittola, T., Huttunen, H., & Virtanen, T. (2015). Polyphonic sound event detection using multi label deep neural networks. *2015 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–7). Killarney, Ireland: IEEE.

Cakir, E., Ozan, E. C., & Virtanen, T. (2016). Filterbank learning for deep neural network based polyphonic sound event detection. *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 3399–3406). Vancouver, BC: IEEE.

Chollet, F. (2018). *Deep learning with Python*. Manning.

Chu, S., Narayanan, S., Kuo, C.-c., & Mataric, M. (2006). Where am I? Scene Recognition for Mobile Robots using Audio Features. *2006 IEEE International Conference on Multimedia and Expo* (pp. 885–888). Toronto, ON: IEEE.

Chun, P., Yamane, T., & Maemura, Y. (2022). A deep learning-based image captioning method to automatically generate comprehensive explanations of bridge damage. *Computer-Aided Civil and Infrastructure Engineering*, *37*(11), 1387–1401.

Couvreur, C., Fontaine, V., Gaunard, P., & Mubikangiey, C. G. (1998). Automatic classification of environmental noise events by hidden Markov models. *Applied Acoustics*, *54*(3), 187–206.

Dozat, T. (2016). Incorporating Nesterov momentum into Adam. *2016 International Conference on Learning Representations (ICLR)-Workshop track*, (pp. 1–4). San Juan, Puerto Rico.

Farhadi, S., Afzal, P., Boveiri Konari, M., Daneshvar Saein, L., & Sadeghi, B. (2022). Combination of machine learning algorithms with concentration-area fractal method for soil geochemical anomaly detection in sediment-hosted Irankuh Pb-Zn deposit, Central Iran. *Minerals*, *12*, 6–689.

Gao, Y., Kong, B., & Mosalam, K. M. (2019). Deep leaf-bootstrapping generative adversarial network for structural image data augmentation. *Computer-Aided Civil and Infrastructure Engineering*, *34*(9), 755–773.

Geiger, J. T., & Helwani, K. (2015). Improving event detection for audio surveillance using Gabor filterbank features. *2015 23rd European Signal Processing Conference (EUSIPCO)* (pp. 714–718). Nice: IEEE.

Giglioni, V., Venanzi, I., Poggioni, V., Milani, A., & Ubertini, F. (2022). Autoencoders for unsupervised real-time bridge health assessment. *Computer-Aided Civil and Infrastructure Engineering*, *38*(8), 959–974.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* (pp. 249–256). JMLR Workshop and Conference Proceedings.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.

Grosse, C., & Ohtsu, M. (2008). *Acoustic emission testing*. Springer.

Hampshire, T. A., & Adeli, H. (2000). Monitoring the behavior of steel structures using distributed optical fiber sensors. *Journal of Constructional Steel Research*, *53*(3), 267–281.

Heittola, T., Mesaros, A., Eronen, A., & Virtanen, T. (2013). Context-dependent sound event detection. *EURASIP Journal on Audio, Speech, & Music Processing*, *1*, 1–13.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv:1207.0580.

Ioffe, S., & Szegedy, C. (2015). *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. Proceedings of the 2015 32nd International Conference on Machine Learning (pp. 448–456), 37, PLMR. Lille, France.

Ito, A., Aiba, A., Ito, M., & Makino, S. (2009). Detection of abnormal sound using multi-stage GMM for surveillance microphone. *2009 Fifth International Conference on Information Assurance and Security* (pp. 733–736). Xi'An China: IEEE.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, & prospects. *Science*, *349*(6245), 255–260.

Khedmatgozar Dolati, S. S., Malla, P., Ortiz, J. D., Mehrabi, A., & Nanni, A. (2023). Identifying NDT methods for damage detection in concrete elements reinforced or strengthened with FRP. *Engineering Structures*, *287*, 116155.

Laxman, K. C., Ross, A., Ai, L., Henderson, A., Elbatanouny, E., Bayat, M., & Ziehl, P. (2023). Determination of vehicle loads on bridges by acoustic emission and an improved ensemble artificial neural network. *Construction and Building Materials*, *364*, 129844.

Lin, Y., Nie, Z., & Ma, H. (2022). Dynamics-based cross-domain structural damage detection through deep transfer learning. *Computer-Aided Civil and Infrastructure Engineering*, *37*(1), 24–54.

Logan, B. (2000). Mel frequency cepstral coefficients for music modeling. *Proceedings of the 1st International Symposium on Music Information Retrieval*, *270*, 1–11.

Luo, H., & Paal, S. G. (2023). A data-free, support vector machine-based physics-driven estimator for dynamic response computation. *Computer-Aided Civil and Infrastructure Engineering*, *38*(1), 26–48.

MacKay, D. J. C. (2019). *Information theory, inference, & learning algorithms* (22nd printing ed.). Cambridge University Press.

Matthews, B. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, *405*(2), 442–451.

Mesaros, A., Heittola, T., Virtanen, T., & Plumbley, M. D. (2021). Sound event detection: A tutorial. *IEEE Signal Processing Magazine*, *38*(5), 67–83.

Murphy, K. P. (2013). *Machine learning: A probabilistic perspective*. MIT Press.

Nair, A., & Cai, C. (2010). Acoustic emission monitoring of bridges: Review and case studies. *Engineering Structures*, *32*(6), 1704–1714.

Nikolenko, S. I. (2021). *Synthetic data for deep learning* (Vol. 174). Springer International Publishing.

Pan, Y., & Zhang, L. (2022). Dual attention deep learning network for automatic steel surface defect segmentation. *Computer-Aided Civil and Infrastructure Engineering*, *37*(11), 1468–1487.

Rafiei, M. H., & Adeli, H. (2017). A new neural dynamic classification algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(12), 3074–3083.

Rafiei, M. H., Gauthier, L. V., Adeli, H., & Takabi, D. (2022). Self-supervised learning for electroencephalography. *IEEE Transactions on Neural Networks and Learning Systems*. https://doi.org/10.1109/TNNLS.2022.3190448

Rao, K. S., & Manjunath, K. E. (2017). *Speech recognition using articulatory and excitation source features*. Springer.

RILEM Technical Committee (2010). Recommendation of RILEM TC 212-ACD: Acoustic emission and related NDE techniques for crack detection and damage evaluation in concrete: Test method for classification of active cracks in concrete structures by acoustic emission. *Materials and Structures*, *43*(9), 1187–1189.

Scruby, C. B. (1987). An introduction to acoustic emission. *Journal of Physics E: Scientific Instruments*, *20*(8), 946–953.

Sigtia, S., Stark, A. M., Krstulovic, S., & Plumbley, M. D. (2016). Automatic environmental sound recognition: Performance versus computational cost. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *24*(11), 2096–2107.

Sirca Jr, G. F., & Adeli, H. (2018). Infrared thermography for detecting defects in concrete structures. *Journal of Civil Engineering and Management*, *24*(7), 508–515.

Stevens, S. S., & Volkmann, J. (1940). The relation of pitch to frequency: A revised scale. *The American Journal of Psychology*, *53*, 3–329.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, *8*(3), 185–190.

Takahashi, N., Gygli, M., Pfister, B., & Van Gool, L. (2016). Deep convolutional neural networks and data augmentation for acoustic event detection. arXiv:1604.07160.

Wang, G., Jia, Q.-S., Zhou, M., Bi, J., Qiao, J., & Abusorrah, A. (2022). Artificial neural networks for water quality soft-sensing in wastewater treatment: A review. *Artificial Intelligence Review*, *55*(1), 565–587.

Yuyama, S., Okamoto, T., Shigeishi, M., & Ohtsu, M. (1995). Quantitative evaluation and visualization of cracking process in reinforced concrete by a moment tensor analysis of acoustic emission. *Materials Evaluation*, *53*, 6.

Yuyama, S., Yokoyama, K., Niitani, K., Ohtsu, M., & Uomoto, T. (2007). Detection and evaluation of failures in high-strength tendon of prestressed concrete bridges by acoustic emission. *Construction and Building Materials*, *21*(3), 491–500.

Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2021). Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, *64*(3), 107–115.

Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2017). Mixup: Beyond empirical risk minimization. arXiv:1710.09412.

Zhang, Y., Macdonald, J. H., Liu, S., & Harper, P. W. (2022). Damage detection of nonlinear structures using probability density ratio estimation. *Computer-Aided Civil and Infrastructure Engineering*, *37*(7), 878–893.

Zheng, Y., Gao, Y., Lu, S., & Mosalam, K. M. (2022). Multistage semisupervised active learning framework for crack identification, segmentation, and measurement of bridges. *Computer-Aided Civil and Infrastructure Engineering*, *37*(9), 1089–1108.

Zhutovsky, S., & Douglas Hooton, R. (2017). Experimental study on physical sulfate salt attack. *Materials and Structures*, *50*, 1–54.

Zou, D., Zhang, M., Bai, Z., Liu, T., Zhou, A., Wang, X., Cui, W., & Zhang, S. (2022). Multicategory damage detection and safety assessment of post-earthquake reinforced concrete structures using deep learning. *Computer-Aided Civil and Infrastructure Engineering*, *37*(9), 1188–1204.

How to cite this article: Farhadi, S., Corrado, M., Borla, O., & Ventura, G. (2024). Prestressing wire breakage monitoring using sound event detection. *Computer-Aided Civil and Infrastructure Engineering*, *39,* 186–202. https://doi.org/10.1111/mice.13079