POLITECNICO DI TORINO
Repository ISTITUZIONALE

Acoustic emission onset time detection for structural monitoring with U-Net neural network architecture

Original
Acoustic emission onset time detection for structural monitoring with U-Net neural network architecture / Melchiorre, Jonathan; D'Amato, Leo; Agostini, Federico; Rizzo, Antonino Maria. - In: DEVELOPMENTS IN THE BUILT ENVIRONMENT. - ISSN 2666-1659. - ELETTRONICO. - 18:(2024), pp. 1-13. [10.1016/j.dibe.2024.100449]

Availability:
This version is available at: 11583/2988403 since: 2024-05-10T11:28:40Z

Publisher:
Elsevier

Published
DOI:10.1016/j.dibe.2024.100449

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

30 May 2026

Developments in the Built Environment 18 (2024) 100449

Contents lists available at ScienceDirect

Developments in the Built Environment

journal homepage: www.sciencedirect.com/journal/developments-in-the-built-environment

Acoustic emission onset time detection for structural monitoring with U-Net neural network architecture

Jonathan Melchiorre [a,*], Leo D'Amato [b], Federico Agostini [c], Antonino Maria Rizzo [d]

[a] Politecnico di Torino, DISEG, Department of Structural, Geotechnical and Building Engineering, Corso Duca Degli Abruzzi, 24, Turin, 10129, Italy
[b] Politecnico di Torino, DAUIN, Department of Control and Computer Engineering, Corso Castelfidardo, 34/d, Turin, 10129, Italy
[c] University of Padua, Department of Physics and Astronomy, Via F. Marzolo, 8, Padua, 35131, Italy
[d] Politecnico di Milano, Department of Electronics, Information Science and Bioengineering, Via Ponzio 34/5, Milano, 20133, Italy

ARTICLE INFO

Keywords:
Acoustic emissions
Onset time detection
Crack localization
Artificial neural network
Segmentation

ABSTRACT

Acoustic Emission (AE) is a non-destructive structural health monitoring technique, which studies elastic waves emitted during crack formation. Utilizing piezoelectric sensors, these waves are converted into electrical signals for subsequent analysis, offering insights into crack propagation and structural durability. This study focuses on the identification of AE signal onset times, crucial for determining crack locations. Conventional methods often encounter challenges with background noise, prompting the need for innovative approaches. Leveraging a U-Net neural network, specialized in segmentation tasks, onset time identification is approached as a one-dimensional segmentation challenge. Through training and testing on Pencil Lead Break (PLB) test data, commonly used in AE evaluations, the effectiveness of the method is demonstrated even with continuous signals, suggesting potential applicability in real-time monitoring.

1. Introduction and literature review

In recent years, rapid advancements in artificial intelligence (AI) have revolutionized various fields (Wang and Siau, 2019), with deep learning emerging as a powerful paradigm for solving complex problems (Bengio et al., 2017). This paradigm shift has been particularly evident in the engineering domain (Tapeh and Naser, 2023; Voulodimos et al., 2018a; Lee et al., 2018), where the introduction of AI techniques has resulted in innovative solutions and crucial insights.

The increasing complexity of engineering problems, combined with the exponential growth of data, has prompted the use of AI methodologies to extract meaningful patterns and make data-driven choices. Among these methodologies, deep learning techniques have garnered significant attention due to their ability to automatically learn hierarchical representations from data, resulting in unparalleled performance in a variety of tasks. Deep learning, a subset of machine learning inspired by the structure and function of the human brain (Mitchell, 1997), has displayed remarkable proficiency in handling complex engineering problems. Deep neural networks are a versatile tool for engineers looking for innovative solutions because of their ability to automatically extract intricate features from large datasets (Melchiorre et al., 2022) and their adaptability to different domains. Artificial intelligence techniques have found applications in a wide range of fields, including pattern recognition (Bishop and Nasrabadi, 2006), material property prediction (Ciaburro and Iannace, 2021), automatic recognition of acoustic sources (Ciaburro, 2020), computer vision (Voulodimos et al., 2018b), and others.

This paper explores a specific application within this paradigm, focusing on deep learning model training for the critical task of detecting onset time in Acoustic Emission (AE) signals.

Acoustic Emission (Scruby, 1987) is a technique employed in structural health monitoring (SHM) tasks. SHM approaches (Sohn et al., 2003; Pasca et al., 2022) are extensively used because they are successful at monitoring historical structures (Manuello et al., 2024), architectural heritage (Farrar and Worden, 2007) and infrastructures (Rosso et al., 2023a; Marasco et al., 2022; Melchiorre et al., 2023a). This enables comprehension of structural dynamics, allowing for prompt interventions with the aim of extending the service life of structures. Among Structural Health Monitoring (SHM) techniques, Acoustic Emission stands out as a valuable tool due to its passive, non-destructive nature (Lacidogna et al., 2015). These characteristics eliminate the need to damage or excite the monitored structure, making it suitable for continuous monitoring of existing structures (Niccolini et al., 2011; Manuello et al., 2019a, 2019b; Lacidogna et al., 2015; Manuello Bertetto

* Corresponding author.
E-mail address: jonathan.melchiorre@polito.it (J. Melchiorre).

https://doi.org/10.1016/j.dibe.2024.100449
Received 15 February 2024; Received in revised form 17 April 2024; Accepted 27 April 2024
Available online 3 May 2024
2666-1659/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

et al., 2020, 2023; Carpinteri et al., 2013; Rosso et al., 2023b). The implementation of these methods includes the use of piezoelectric sensors capable of recording the propagation of transient ultrasonic waves. These waves originate from the sudden release of elastic energy occurring during the formation of cracks. These waves typically have a frequency spectrum ranging from 1 kHz to 10 GHz (Gorman, 1991).

The piezoelectric phenomenon (Arnau et al., 2004) allows the ultrasonic waves to be converted by the sensors into an electric voltage that can be digitized and analyzed (Aggelis, 2011). The examination and analysis of the recorded signals allow deriving indirect insights into the nature of the formation and evolution of cracking patterns. In recent years, numerous techniques have been developed to localize, characterize, and quantify damage based on the study of acoustic emission signals (Ohtsu, 1987; Carpinteri et al., 2007; Ohtsu et al., 1998).

In general, the application of Acoustic Emission (AE) is associated with two primary tasks: the classification of crack typology and the localization of the crack source. In this study, the emphasis is on the second task, as knowing the location of the crack source is crucial for comprehending the causes of damage and facilitates prompt maintenance interventions. The onset time of an AE signal is defined as the first time the elastic wave reaches the piezoelectric sensors (Carpinteri et al., 2006). The accuracy in determining the onset time in AE signal is a critical aspect of damage localization because it directly impacts the precision of locating the crack event. As a result, various techniques for automatic detection of onset time have been developed over the years (Bai et al., 2017; Kurz et al., 2005).

The threshold methods constitute the first category of techniques, relying on static or dynamic thresholds to discern when the signal transitions from background noise to actual Acoustic Emission (AE). Among these, the amplitude threshold method is the most common (Eaton et al., 2012; Sedlak et al., 2013), defining the onset time as the moment when the signal amplitude surpasses the predetermined threshold value (Rocchi et al., 2019). The primary challenge in using this method is determining the appropriate threshold for background noise. Selecting a threshold value that is too low may result in premature triggering by the preceding background noise, while selecting a value that is too high may result in missing the actual onset time of the signal. The optimization of the detection threshold value may be achieved for specific signal-to-noise ratio (SNR) values if known in advance. However, in real-world applications, a broad range of SNRs can be encountered, introducing vulnerability to measurement errors. Enhancements to the threshold approach include the use of complex dynamic threshold values, where the threshold is updated based on the average acoustic noise amplitude. The STA/LTA technique (STA - Short Term Average, LTA - Long Term Average) is based on this approach (Baer and Kradolfer, 1987).

Other methods for determining the onset of AE signals have been developed in both the time as well as the time-frequency domains.

In the time domain, Akaike (1974) introduced a statistical method for identifying the transition point in a time-series between noise and a coherent signal in the time domain. This method is commonly referred to as the AIC picker as it relies on the use of the Akaike Information Criterion (AIC) to autoregressive models. The main idea is to allow the proper division of acoustic emission signals into two stationary datasets—before and after the onset time (Kurz et al., 2005). To enhance the accuracy of this method, an improved AIC procedure was introduced (Carpinteri et al., 2012). This technique relies on estimating the accuracy of AE signals through the second derivative of the AIC function and a parameter associated with the propagation velocity of the elastic waves.

To deal with signals with low signal-to-noise (S/N) ratios, a method relying on the variation in fractal dimension along the trace has been introduced (Boschetti et al., 1996). This fractal-based algorithm has demonstrated accuracy even in the presence of substantial noise but is considered to be considerably slower than other methods and is not well-suited for real-time applications.

The Hinkley criterion (Hinkley, 1971) is another time-domain statistical method that calculates the partial energy of a time-series for all samples of the signal. In this case, the global minimum of the partial energy function serves as an indicator of signal onset.

In the time-frequency domain, certain methods involve the cross-correlation of the AE signal with a short Gaussian pulse at a specific frequency (Ziola and Gorman, 1991). The concept is that the onset time at a particular frequency is indicated by the peak of the correlation function.

A similar method (Ciampa and Meo, 2010) for obtaining the time-frequency response has been developed, which employs the Continuous Wavelet Transform (CWT) (Sadowsky, 1996). In this case, the reference for the Time Difference of Arrival measurements was determined based on the time corresponding to the maximum of the squared modulus coefficients of the CWT, deviating from the conventional approach of using the signal onset.

Traditional signal processing techniques, as discussed, have found widespread application for automatic onset time detection. However, despite their proven efficacy, these methods exhibit certain drawbacks when applied in practical, on-field scenarios. For instance, the threshold-dependent method shows instability in the presence of noise, particularly when the frequencies associated with background noise closely resemble those of the signals, a common situation in concrete structures. Additionally, other methodologies often rely on some form of predetermined threshold determined through trial and error or exhibit inherent instability. Moreover, some methods necessitate significant computational time, making them unsuitable for real-time applications.

Therefore, in recent years, novel methodologies have been proposed, leveraging on the advent of Artificial Intelligence and Machine Learning techniques.

An early example of employing machine learning techniques for localizing the source of acoustic emission signals was pioneered by Emanian et al. (Emamian et al., 2003). The method involves eliminating background noise in AE signals through a combination of covariance analysis (Bilgen and Insana, 1998), Principal Component Analysis (PCA) (Wold et al., 1987), and differential time delay estimates (Weiss and Weinstein, 1983). Subsequently, a self-organizing map (SOM) neural network (Kohonen, 1990) is employed to differentiate between the noise and AE signals.

In 2003, Ince et al. (2010) proposed a method based on the Support Vector Machine (SVM) algorithm (Suthaharan and Suthaharan, 2016). In this case, hierarchical clustering and SVM are introduced for clustering AE signals and detecting P-waves for microcrack locations in the presence of noise. Pairwise correlation analysis (Thompson et al., 2012) is used to identify clusters of AE events. Afterwards the identification of clusters, an averaging step is performed to obtain “super” AE with an improved Signal-to-Noise Ratio (SNR). The characteristic features in the time and frequency domains are extracted from the data by employing autoregressive modeling (Sodsri, 2003), wavelet packets (WP) (Maradei et al., 2003), and discrete Fourier transform (Sundararajan, 2001). Finally, These features are utilized to train SVM classifiers with probabilistic outputs to obtain the onset time of the signals.

In 2020, Zhang et al. (2020) conducted a study comparing three machine learning classifiers, combined with continuous wavelet transform (CWT) preprocessing, to achieve high-precision onset time detection. The comparison included an Extreme Learning Machine (ELM) classifier (Wang et al., 2022), a decision tree classification model (DTC) (Priyanka, 2020), an ensemble tree model (RFC) (Weinberg and Last, 2019), and a deep belief network (DBN) classifier (Hua et al., 2015).

In 2020, Chen et al. (2020) proposed using deep learning models, specifically Convolutional Neural Network (CNN) (Gu et al., 2018), for onset time detection. The method employs the short-time Fourier transform as a feature extraction method to provide the input parameters to the CNN classifier.

In 2021, Jierula et al. (2021) adopted a back propagation neural network model (Buscema, 1998) for detecting damage locations in pile

foundations using acoustic emission signals.

In 2022, Zonzini et al. (2022) demonstrated the superior accuracy of deep learning models compared to traditional techniques for onset time detection in highly noisy scenarios. Specifically, they compared the accuracy of a Convolutional Neural Network and a capsule neural network (Patrick et al., 2022) with the classical Akaike Information Criterion (AIC). The results indicate that the deep learning models achieved a 10 × improvement in accuracy.

In 2023, Melchiorre et al., 2023b, 2023c demonstrated the robustness of the deep learning models in the onset time detection. A Convolutional Neural Network (CNN) for Sound Event Detection (SED) (Mesaros et al., 2021) was trained on normalized seismic accelerograms and subsequently tested it on Acoustic Emission (AE) signals. The models demonstrated high accuracy in identifying the onset time.

The applications of machine learning and deep learning techniques mentioned above have demonstrated the effectiveness of using new methods in the field of noise emissions. These techniques have proven to be very robust and capable of increasing accuracy in determining the onset time in acoustic emission signals. However, research in this field is still ongoing as the methodologies shown still have limitations related to the use of computationally and storage-intensive algorithms or the need for data preprocessing techniques to improve the SNR. Furthermore, these techniques operate on signal segments, necessitating the use of preprocessing techniques to define signal windows suitable for classification through machine learning algorithms.

This paper introduces a novel deep learning model for detecting onset times in acoustic emission signals. The method utilizes a U-Net neural network (Ronneberger et al., 2015), specifically designed for segmentation tasks. The onset time detection problem is transformed into a one-dimensional segmentation task, with the model trained to classify each point in the acoustic emission signals. Each point in the time-series is assigned a probability of belonging to the signal or background noise. To minimize false positives, a rolling average probability is implemented, smoothing the probability function along the time-series and mitigating issues resulting from signal discontinuities.

Various dimensions of the proposed architecture were tested to identify a model achieving accurate onset time detection while being computationally and storage-efficient. This resulted in a lighter model compared to the more commonly used convolutional neural networks in the acoustic emissions field. Furthermore, the method is designed to work directly on continuous signals without the need for dataset preprocessing techniques. The objective is to create a technique applicable directly in the field with the ambition of real-time onset time identification.

The effectiveness of the proposed method was assessed using a dataset derived from a Pencil Lead Break (PLB) test (Madarshahian et al., 2019a). The PLB test is a test specifically employed to replicate acoustic emission signals for the calibration of instrumentation and methodologies. The core concept involves fracturing a pencil lead against a concrete block. This fracture produces an elastic wave closely resembling those generated by cracks. The elastic wave travels from the pencil to the concrete block and can be recorded by piezoelectric transducers. A significant advantage of this approach is the ability to ascertain the location of the fracture. This allows the assessment of the precision of both the instrumentation and the techniques employed for crack localization.

2. Methodology

2.1. The importance of the onset time detection

The acoustic emissions are transient stress waves generated by the release of energy within a material, often associated with the initiation or propagation of cracks, fractures, or other structural defects.

The analysis of acoustic emissions signals is widely employed in structural monitoring as it enables the identification of defects and damage within the structures. Generally, this method is utilized for monitoring, localizing, and quantifying damage in a structure during its operational conditions. This allows for the quantification of the energy released during the propagation of fractures within the structures. Moreover, the analysis of the obtained data enables the assessment of the durability performance of the structures. The monitoring system employed for the acoustic emission technique consists of various piezoelectric sensors that digitize the elastic wave emitted during crack propagation. This system is typically installed on the structure during its operational life. It records the AE signals emitted by fractures resulting from the stress to which the structure is subjected during its service life. In this way, the monitoring does not necessitate active loading of the structure, enabling passive and non-destructive monitoring of the structure itself.

Among the various parameters that can be obtained from the analysis of AE signals, one of the most important is the onset time, because the precise identification of onset time in acoustic emission signals is crucial for accurately localizing cracks.

In a simplified scenario, the elastic wave travels directly from the emission point to the piezoelectric sensors. This implies that the wave path can be represented as a direct line connecting the point of the crack occurrence and the sensors. The shortest wave path model is a common simplification in the field of acoustic emissions (Carpinteri et al., 2012; Madarshahian et al., 2019b). In particular, this assumption can be adopted for the onset time identification methods. This is because, although the signals of both P-waves (longitudinal waves) and S-waves (shear waves) could be used for crack characterization, the onset time identification is generally performed only on P-waves signals. This choice is motivated by their reduced vulnerability to disturbances from multiple side reflections, structural noise, and sensor response.

Under the assumption of the shortest path model, the distance d_{0-A} between the crack source at $S_0 = (x_0, y_0, z_0)$ and the location of the piezoelectric sensor A, $S_A = (x_A, y_A, z_A)$ can be calculated using the Equation (1).

$$d_{0-A} = \sqrt{(x_{0(1)} - x_A)^2 + (y_0 - y_A)^2 + (z_0 - z_A)^2} \tag{1}$$

If the medium is assumed to be homogeneous, the wave speed c can be considered constant, and the time it takes for the elastic wave to travel from the crack source to the piezoelectric sensor can be obtained using Equation (2).

$$T_A = \frac{d_{0-A}}{c} = \frac{\sqrt{(x_0 - x_A)^2 + (y_0 - y_A)^2 + (z_0 - z_A)^2}}{c} \tag{2}$$

In general, the source location (x_0, y_0, z_0), the absolute time of the crack event t_0 and the wave speed c are unknown. This implies that Equation (2) cannot be solved directly. However, by analyzing the AE signals, it is possible to retrieve the relative arrival times at each transducer Δt_A. Thus, Equation (2) can be rewritten by subtracting the wave arrival time from a reference transducer (T_R) from both sides of the equation, as follows:

$$\Delta t_A = T_A - T_R = \frac{\sqrt{(x_0 - x_A)^2 + (y_0 - y_A)^2 + (z_0 - z_A)^2}}{c} - T_R \tag{3}$$

Equation (3) can be reformulated for each sensor. Therefore, with a minimum of 5 sensors, it is possible to establish a system of equations that determines all the unknowns in the problem. Specifically, this implies that by obtaining the onset time of the AE signal recorded by at least 5 different sensors, it becomes feasible to calculate the precise location of the crack.

2.2. Deep learning architecture

Section 2.1 outlines the reasons why the identification of onset time in Acoustic Emission signals is crucial for crack localization. While this

task is relatively straightforward for expert users to address manually, the challenge arises with the automatic identification of the onset time, which is essential for enabling automated and real-time structural monitoring. Moreover, in practical monitoring scenarios, it is typical to record thousands of signals, rendering manual onset time detection virtually impossible.

As discussed in the literature review provided in Section 1, numerous methods have been developed over time for this task. In particular, various solutions involve the application of machine learning methods for onset time identification. In this paper, an innovative deep learning-based method is introduced for onset time identification. The approach addresses a one-dimensional binary classification problem, determining whether each point in the recorded signal is related to background noise or constitutes an integral part of the AE signal.

The methodology relies on a neural architecture known as U-Net (Ronneberger et al., 2015). This architecture represents a specialized type of Convolutional Neural Network (CNN) tailored for addressing segmentation tasks, particularly in the realm of image processing. The U-Net model finds its primary application in image segmentation, with numerous studies in the literature showcasing its efficacy, particularly within the medical domain (Du et al., 2020; Siddique et al., 2021).

The U-Net architecture is structured as an encoder-decoder model (see Fig. 1), facilitating the capture of both local and global context information effectively. It comprises two primary components: the contracting path (encoder) and the expansive path (decoder) (Long et al., 2015).

In the contracting path or encoder, a series of convolutional and pooling operations are applied to progressively decrease the spatial dimensions of the input image while increasing the number of feature channels. This process aids in extracting high-level features and forming a compressed representation of the input. The encoder consists of multiple repetitive blocks, each containing consecutive convolutional layers with ReLU activation, followed by max-pooling layers. These blocks aim to capture increasingly abstract features while reducing spatial resolution.

Conversely, the expansive path, or decoder, employs a sequence of up-convolutional and concatenation operations to gradually restore the spatial information lost during encoding. This path enables the network to generate a segmentation map matching the spatial dimensions of the input image. The decoder comprises multiple repetitive blocks, featuring up-convolutional layers to boost spatial resolution, followed by concatenation with corresponding feature maps from the contracting path. This combination helps merge low-level and high-level features, critical for precise localization.

The adopted U-Net encoder comprises N convolutional blocks, where experiments were conducted for $N \in$ (Wang and Siau, 2019; Bengio et al., 2017; Tapeh and Naser, 2023; Voulodimos et al., 2018a). Each block contains two 1D convolutional layers, each followed by a GELU activation function. Similarly, the decoder mirrors the encoder but employs upsampling to enhance temporal resolution. Finally, the output layer consists of a 1D convolutional layer followed by a softmax activation function.

A notable feature of the U-Net architecture is the incorporation of skip connections, linking feature maps from the contracting path directly to corresponding blocks in the expansive path. This design enables the decoder to access and utilize multi-scale information from earlier network stages, enhancing segmentation accuracy.

Towards the end of the expansive path, a 1×1 convolutional layer with softmax activation is applied to generate a pixel-wise probability map. Each pixel in the output map indicates the probability of belonging to a specific class. During training, the network is optimized using appropriate loss functions like dice loss or cross-entropy loss, aiming to minimize the disparity between predicted segmentation maps and ground truth.

In summary, the U-Net architecture offers a robust framework for segmentation tasks, leveraging the advantages of both global and local context information through its encoder-decoder structure and skip connections.

2.3. Dataset

The dataset utilized to train the U-Net neural network is obtained by (Madarshahian et al., 2019a). The dataset discussed in the paper originates from artificial acoustic emission (AE) tests conducted on a concrete block. These tests utilized pencil lead breaks (PLBs) to generate elastic waves for data collection.

The data were collected from a concrete block specimen measuring 30.5 cm $\times$ 31.0 cm $\times$ 112 cm composed of Portland cement, coarse aggregate, natural sand, and a sodium hydroxide solution to achieve a water-to-cement ratio of 50%.

Ten piezoelectric AE sensors, operating within a frequency range of 200–850 kHz, were then affixed to the specimen using epoxy and stainless steel holders.

The sensors were connected to a 24-channel Express-8 AE data acquisition system data acquisition system with a threshold of 31 dB manufactured by MISTRAS Group (MISTRAS Group, 1978). This setup facilitated the recording of stress waves emitted by the PLBs. PLBs, each with a diameter of 0.3 mm, were subsequently broken ten times at three predetermined locations. The resulting stress wave signals captured by the sensors were then amplified, digitized, and recorded by the data acquisition system. The recorded signals have been disclosed in CSV files. Each file contains metadata such as the test date, test time, sample interval, signal unit, number of data points, and channel numbers.

Fig. 2 presents examples of signals from the dataset depicted both as time-series and as spectrograms in the frequency domain. The spectrograms are generated by applying a Fourier transform to the signals. It is apparent from the figure that distinguishing the point separating the background noise from the acoustic emission signal is not straightforward in the spectrograms. This indicates that the background noise often shares similar frequencies with the signal, particularly in concrete structures. Consequently, filtering signals to eliminate or mitigate background noise is not always feasible. As a result, the accuracy of many automatic onset time identification techniques, such as threshold-based methods, is diminished.

The dataset consists of 98 time-series sampled at a rate of 1000 kHz, corresponding to a sampling period of $T = 1$ μs. Each AE signal consists

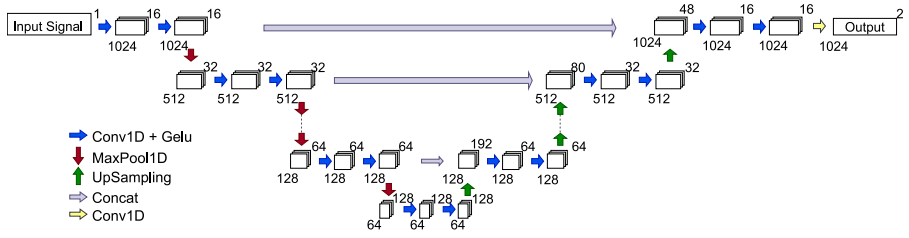

Fig. 1. Architecture of the U-Net model implemented in the analysis of AE signals. The output is the probability for each time stamp to be *signal* rather than *background* noise.

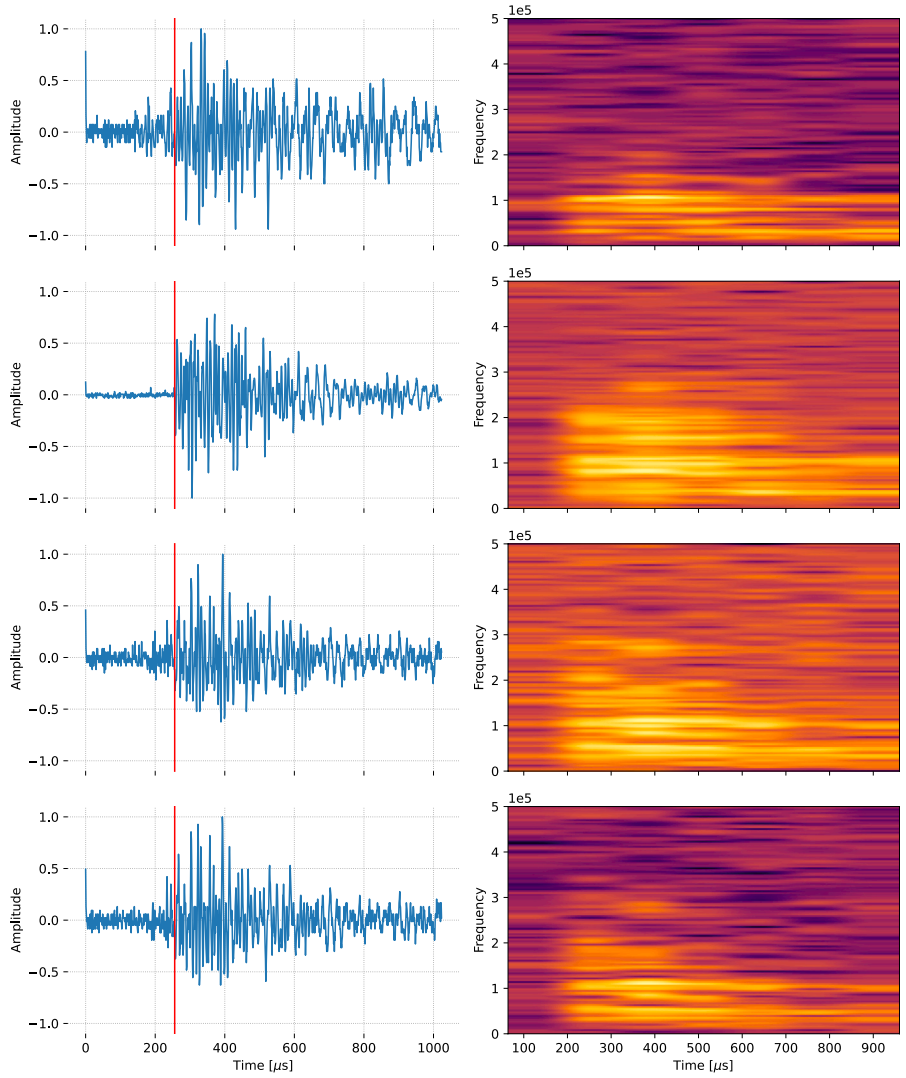

Fig. 2. Illustration depicting examples of signals from the utilized dataset, presented in both time-series format (left) and as spectrograms (right). The amplitudes are normalized, while the frequencies are denoted in Hertz [*Hz*].

of 1024 time samples, with the onset time event preceded by a constant pre-triggering period of 256 μs This setup facilitates straightforward and automated labelling of the signals. In particular, label 0 has been assigned to the background noise, and label 1 is assigned to the actual signal.

In this study, the objective is to develop an automated methodology for identifying the onset time in acoustic emission (AE) signals, applicable in real-time scenarios. Unlike traditional approaches involving preprocessing techniques to segregate AE signals into distinct time-series, a different method is adopted here. The signals in the training dataset for the neural network are amalgamated into a single time-series, achieved by randomly shuffling the signals.

To ensure that the network learns the features of acoustic emissions rather than the symmetry of the input data, the dataset is processed in a specific manner. This involves ensuring that the onset time does not consistently occur at a fixed time step, such as 256. Instead, all the time-series are concatenated to form a single long signal, which is then segmented with a stride of multiple windows of varying lengths. Each window becomes a new time-series representing a data point in the new dataset. Crucially, each window may or may not contain an onset time, and if it does, the onset time can occur at any time step within the window.

Additionally, in the initial part of each signal containing background noise, a random segment was trimmed to create a time-series with non-uniformly spaced acoustic emission signals. Subsequently, the resultant signal was partitioned into a 75% training set and a 25% validation set.

2.4. Neural network training and data segmentation

This study attempted to reduce computing costs and streamline the approach in order to create the basis for the development of a real-time monitoring system. Hence, it was chosen to utilize the time-series data directly as input for the neural model.

This is in contrast to the approach adopted in many existing works, which relies on spectrogram-derived signal segmentation commonly utilized in convolutional networks for image segmentation (Melchiorre et al., 2023b, 2023c). This choice prevents the need for Fourier transform preprocessing of the signals. Moreover, processing the time-series as one-dimensional input data reduces computational costs significantly when compared to methods that employ two- or three-dimensional input data, such as those leveraging image colors.

The model assigns a probability to each signal point, indicating the likelihood of it belonging to either class. Consequently, every time-step receives a probability score, indicating its probability of being associated with either background noise or the AE signal.

In general, given an input signal $x = (x_1, \ldots x_N)$, the output of the U-

Net neural network yields a vector of probabilities $p = (p_1, \ldots p_N)$. Here, each p_i denotes the probability of x_i being part of either the background noise or the actual AE signal.

To obtain the final label, a threshold on the probability determined by the neural model T is set, and all points with a probability $q_i > T$ are classified as the AE signal, otherwise, they are classified as part of the background noise. The onset time is denoted by the point where there is a transition from background noise to the AE signal, indicated by the probability exceeding the defined threshold.

The neural model has been adapted to process the one-dimensional signal as input and conduct segmentation to classify each point in the signal accordingly. Two distinct classes have been established: class 0, encompassing all points attributed to background noise, and class 1, designated for all points representing the actual AE signal.

The model then generates the probability for each signal point to belong to either class. Consequently, each time-step is assigned a probability indicating its likelihood of being associated with either the background noise or the AE signal.

An illustration of signal classification is depicted in Fig. 3. Here, the segment of the signal highlighted in blue denotes the background noise, while the portion in red signifies the AE signal. The onset time is visually indicated by a vertical green line, marking the transition from background noise to AE signal.

In this study, various neural models were evaluated by adjusting the length of the input sequence and the depth of the U-Net architecture. Specifically, outcomes were compared across three distinct input sequence lengths: 512, 1024, and 2048 samples. Furthermore, the examination involved three different depths of the neural model: 2, 3, and 4. Here, the depth signifies the number of encoding/decoding blocks employed in the architecture.

During the training phase, careful consideration was given to setting the maximum number of epochs, which was established at 200 iterations. This decision aimed to strike a balance between computational efficiency and ensuring adequate model convergence. Additionally, to prevent overfitting and ensure optimal performance, an early stopping mechanism was implemented. This mechanism, based on the loss metric, allowed training to halt if no improvement was observed in the validation set performance over a specified number of epochs, set at 20 in this study.

The choice of loss function is a critical aspect of neural network training, influencing the ability of the model to accurately capture the underlying patterns in the data. In this study, binary cross-entropy was selected as the loss function due to its suitability for binary classification tasks, such as distinguishing between background noise and acoustic emission signals. This decision was supported by its widespread adoption and proven effectiveness in similar applications (Ruby and Yendapalli, 2020; Ho and Wookey, 2019).

Furthermore, the optimization algorithm employed to update the model parameters plays a pivotal role in the training process. The ADAM optimizer, renowned for its efficiency and robustness, was chosen for this task. Configured with a learning rate of 10^{-4}, ADAM dynamically adjusts the learning rate during training, facilitating faster convergence and improved performance. This combination of hyperparameters was aimed at optimizing the training process, enabling the model to effectively learn the underlying patterns in the data.

It is important to note that in this preliminary study, manual exploration of a suitable combination of the hyperparameters of the model was employed, given its validity and widespread adoption (Anitescu et al., 2019).

In Fig. 4, the loss and accuracy of the models across various training epochs are depicted, with blue representing metrics for the test set and red for the validation set. The graphs illustrate the performance of models with different depths and varying input vector lengths.

Early stopping was employed in the majority of models, indicating their susceptibility to overfitting. Specifically, only in the larger models with four encoding/decoding blocks and an input segment length of 2048 samples, the training continued through all 200 epochs. Despite variations in training duration, it is evident that all models ultimately converge to comparable final loss and accuracy values.

Following the model training, it was observed that all classification errors in the output exhibited a consistent pattern. Specifically, there were instances where the model erroneously categorized samples from the background noise as part of the AE signal. However, it was exceedingly rare for the model to make the opposite error, i.e., misclassify samples from the signal as part of the background noise. An illustration of misclassification error is depicted in the initial graph of Fig. 5. Here, three onset time predictions (indicated by red dashed lines) are compared with a single real onset time (depicted by a green solid line).

The misclassification mentioned above leads to the occurrence of spurious onset times in the signal. Specifically, each misclassification error, followed by the correct classification of the subsequent point as background noise, results in the definition of a spurious onset time. Table 1 provides an overview of the number of spurious onset times obtained by the various models, considering both the training and validation sets.

This characteristic tendency is inherent to the approach adopted and had been previously noted in (Melchiorre et al., 2023c). In that instance, adjustments to the threshold T were made to mitigate this issue. Specifically, the probability threshold value required to classify a point as part of the AE signal was increased to values exceeding 90%.

In this study, a novel approach was adopted to mitigate the occurrence of spurious onset times. The proposed method involves smoothing the probability values obtained from the output of the U-Net, thereby reducing the impact of misclassification errors. Misclassifications often lead to localized fluctuations in the probability values of neighboring points, as illustrated in Fig. 5.

The smoothing process entails averaging the probability values over

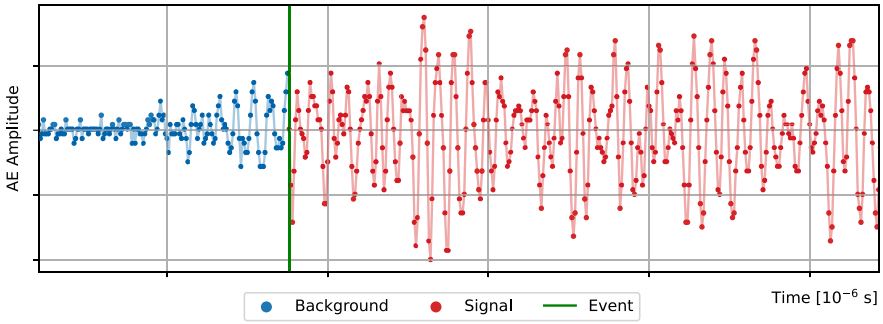

Fig. 3. Each data point within the time-series undergoes classification as either "background" or "signal" by a U-Net model. The onset time is identified when the classification label transitions from "background" to "signal".

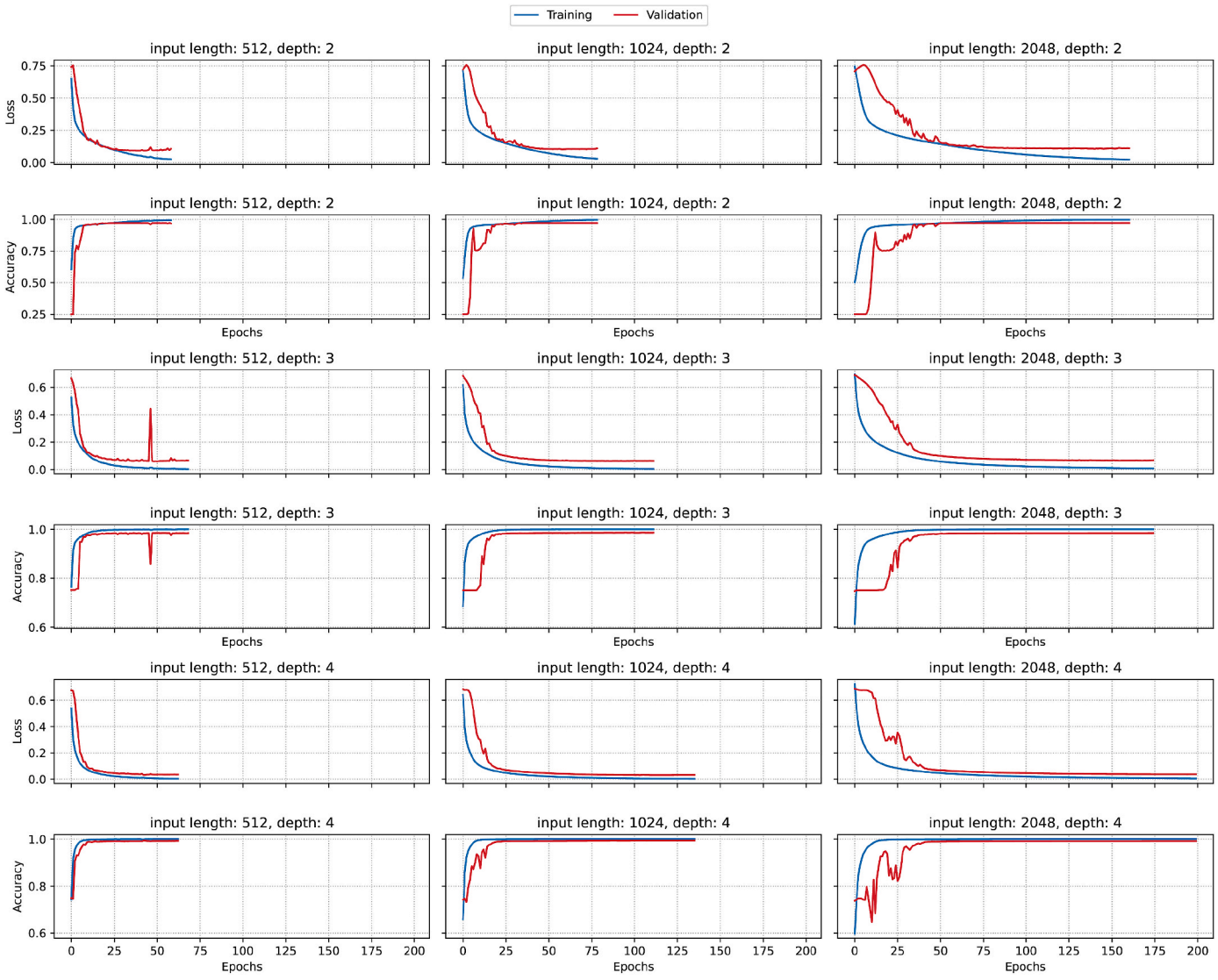

Fig. 4. Training and validation loss and accuracy for the U-Net models with different lengths of the input, and a depth of 4.

a window comprising 50 samples. This rolling averaging procedure is applied along the signal, generating a smoothed probability function (refer to Fig. 5). Subsequently, a threshold of $T = 0.5$ is applied to the smoothed function, serving as the basis for classifying samples at each time step. Unlike the direct classification based on the probabilities provided by the neural model, this approach utilizes the smoothed probabilities derived from the rolling average.

This methodology yields a significant reduction in the occurrence of spurious onset times. Table 1 presents the number of spurious onset times obtained under both the direct classification based on the neural model probabilities and the corrected classification employing the rolling average method.

The adopted method demonstrated significant effectiveness, with instances where it reduced the number of spurious onset times by over 90%.

3. Results

This section presents a comparative analysis of the results achieved with various neural models. The comparison focuses on evaluating each model using classical machine learning metrics.

In this comparison, the assessment of different neural models includes a range of evaluation metrics such as Accuracy, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and F1-score. These metrics are the classical metrics used to evaluate machine learning models and are defined as follows:

$$\text{Accuracy} = \frac{\text{Number of True Predictions}}{\text{Total Number of Predictions}} \tag{4}$$

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\text{True Value}_i - \text{Predicted Value}_i|}{n} \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\text{True Value}_i - \text{Predicted Value}_i)^2}{n}} \tag{6}$$

$$\text{F1} - \text{score} = \frac{1}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \tag{7}$$

Where:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{8}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{9}$$

The various neural models, along with their corresponding

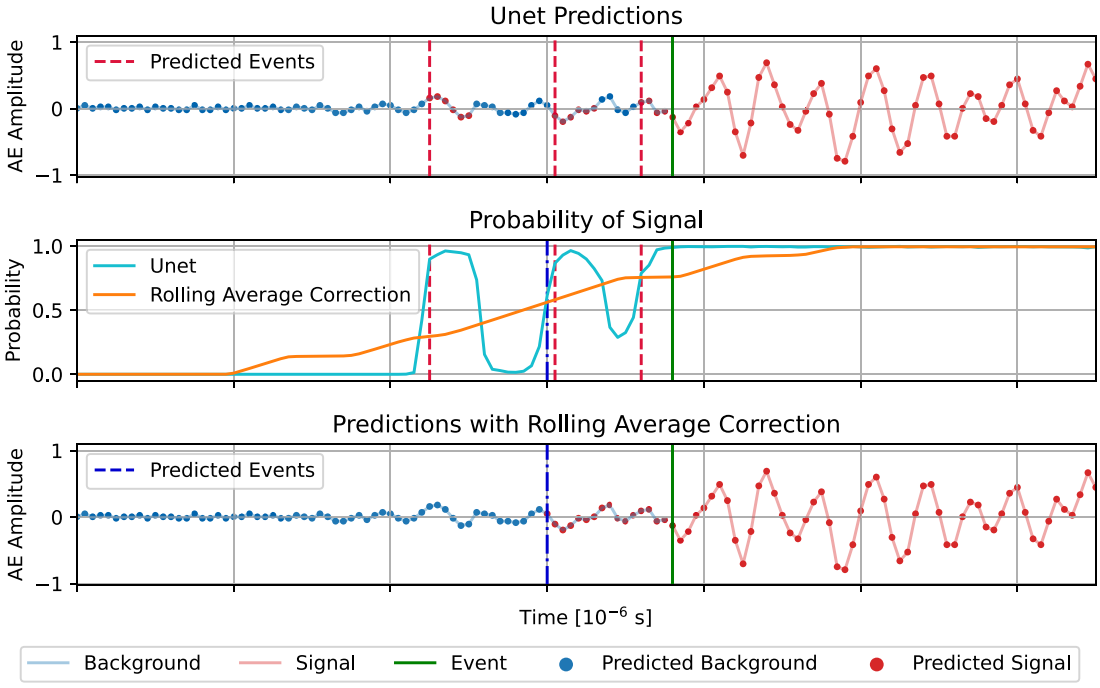

Fig. 5. Effects of smoothed probability with rolling average correction. The top section displays predictions based on model output probabilities. In the middle section, a comparison is shown between U-Net probabilities and the rolling average correction. The bottom section presents predictions using the smoothed probability. Dashed red vertical lines indicate predicted onset times from U-Net probabilities, solid green vertical lines represent actual onset times, and blue dash-dot lines depict onset time predictions after applying Rolling Average correction. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Number of spurious onset time detected by the U-Net outputs before and after applying the rolling average correction.

Model		U-Net Predictions		R.A. Correction	
Seq. Len.	Depth	Train	Validation	Train	Validation
2048	4	0	19	0	1
2048	3	0	57	0	10
2048	2	191	242	9	13
1024	4	0	21	0	2
1024	3	0	56	0	7
1024	2	485	270	32	24
512	4	5	22	2	2
512	3	0	70	0	5
512	2	625	282	37	21

evaluation metrics, are summarized in Table 2. The metrics are presented for both the training and validation sets.

Specifically, Table 2 displays the outcomes obtained by varying the length of the input sequence considered during model training and the depth of the U-Net architecture, ranging from two to four levels. The inclusion of a longer input sequence aims to enable the model to capture the context surrounding the event and improve its ability to identify the event onset time. Additionally, increasing the number of levels and parameters in the ANN allows for a better fit to the dataset, paying attention to avoiding the overfitting phenomenon.

The metrics presented in the table showcase the models effectiveness in accurately classifying signal samples and distinguishing between background noise and AE signal. Notably, all models achieve accuracies exceeding 0.97, with some reaching peaks above 0.9999. Furthermore, the mean absolute error (MAE) values demonstrate consistently low errors across the board, with the best result recorded at 0.006 when evaluated on the test set.

The metrics illustrated in Table 2 underscore remarkably high-performance levels. However, delving deeper into understanding the factors driving such impressive outcomes is imperative. It is crucial to acknowledge that the utilized metrics evaluate the effectiveness of the model in categorizing individual samples within the time-series. To gain a comprehensive understanding of the achieved results, it is vital to consider the distinct structure of the time-series. Initially, the time-series

Table 2
Performance comparison of various U-Net models based on classical machine learning metrics.

Model		Train						Validation					
Seq. Len.	Depth	Loss	Acc	MAE	RMSE	F1	AUC	Loss	Acc	MAE	RMSE	F1	AUC
2048	4	0.006	1.000	0.006	0.007	1.000	1.000	0.037	0.992	0.014	0.084	0.995	0.997
2048	3	0.010	1.000	0.010	0.013	1.000	1.000	0.065	0.985	0.026	0.115	0.990	0.991
2048	2	0.031	0.997	0.028	0.056	0.997	1.000	0.108	0.973	0.052	0.153	0.980	0.986
1024	4	0.004	1.000	0.004	0.005	1.000	1.000	0.031	0.994	0.011	0.072	0.996	0.998
1024	3	0.007	1.000	0.007	0.010	1.000	1.000	0.061	0.986	0.021	0.109	0.991	0.991
1024	2	0.055	0.991	0.047	0.096	0.991	0.999	0.103	0.972	0.062	0.151	0.978	0.990
512	4	0.012	0.999	0.008	0.019	1.000	1.000	0.032	0.993	0.015	0.072	0.996	0.998
512	3	0.008	1.000	0.007	0.013	1.000	1.000	0.059	0.985	0.023	0.112	0.990	0.993
512	2	0.052	0.987	0.040	0.095	0.990	0.999	0.092	0.971	0.054	0.148	0.978	0.993

consistently exhibits the presence of background noise, which is distinguishable from the subsequent emergence of the AE signal that arises when the wave reaches the sensor.

In general, distinguishing between most samples belonging to these two phases along the time-series is not so challenging. However, accurately pinpointing the transition between background noise and the AE signal proves to be more difficult. Consequently, classical machine learning metrics may yield high performance, as the easily discernible initial and end samples heavily influence the evaluation of the neural models. Nevertheless, this evaluation method provides limited insight into the proximity of the predicted onset time to the actual onset time. Given that the transition time represents only a small fraction of the entire time-series, even a significant discrepancy between predicted and actual onset times may still result in high metrics scores.

In this paper, to provide a comprehensive evaluation of the proposed models, a comparison was made based on the time difference $\Delta t[\mu s]$ between the predicted onset time and the actual onset time. This evaluation metric offers more meaningful insights than classical machine learning metrics. The discrepancy between the real and predicted onset times reflects the true error of the proposed neural models. This measure is directly linked to the accuracy of crack location identification. A larger difference between these two quantities indicates a greater error in crack evaluation.

In this task, minimizing Δt is crucial due to the high propagation speed of AE signals. Assuming a propagation velocity of 4000 m/s, typical of wave propagation in concrete, even a small error in identifying the onset time can result in a significant error in locating the crack source.

The average absolute temporal discrepancies Δt between predicted and actual onset times, measured in μs, are presented in Table 3 for all evaluated neural networks. Results are provided for both direct classifications using U-Net-calculated probabilities and the smoothing probability method detailed in Section 2.4.

Employing the corrective approach on model outputs notably enhances the precision of the proposed methodology. The most optimal model achieves an average error of 7 μs, corresponding to an approximate positional error of 2.8 cm, considering an acoustic wave velocity of 4000 m/s. This improvement stems from a reduction in the incidence of false positive onset time events, which inflates the mean discrepancy with genuine onset times.

Fig. 6 illustrates a comparative analysis of all trained models, highlighting the variations in terms of Δt that is possible to achieve through the proposed methodology. The findings indicate that the primary factor influencing the outcomes is the depth of the U-Nets. As expected, heightened depth correlates with a notable enhancement in the accuracy of neural models in precisely identifying the onset time. Conversely, the length of the input signal sequence appears to exert less impact on the outcome.

It is important to highlight that the attained high accuracy was accomplished without resorting to data preprocessing. Moreover, the study did not consider potential enhancements achievable by concurrently integrating multiple signals from piezoelectric sensors. While these techniques hold promise for further enhancing method accuracy, they fall outside the scope of this study. The research focuses on crafting a deep learning approach for real-time onset time identification in continuous signals.

Table 3
Average Δt (10^{-6} s) between predicted and actual onset times, with an estimated positional error assuming an acoustic wave propagating at a speed of 400 m/s.

Model		U-Net Predictions		R.A. Correction		
Seq. Len.	Depth	Train	Validation	Train	Validation	Δx (cm)
2048	4	0	14.63	0	7.03	2.812
2048	3	0	86.27	0	21.34	8.536
2048	2	142.83	201.98	4.74	28.56	11.424
1024	4	0	25.58	0	10.62	4.248
1024	3	0	81.71	0	22.12	8.848
1024	2	211.24	207.67	19.59	35.64	14.256
512	4	5.46	33.53	3.39	9.01	3.604
512	3	0.11	89.39	0.11	16.51	6.604
512	2	242.55	210.33	25.37	36.02	14.408

3.1. Comparison with other acknowledged methodologies

In this section, the proposed method is validated through comparisons with widely acknowledged signal analysis techniques aimed at distinguishing signal from background noise. Specifically, the method is compared with two different techniques: a threshold approach and the Change Point Method (CPM) for Change Detection.

These methods operate under the assumption that the onset time aligns with a notable shift or fluctuation in the signal. However, this assumption is not always valid, and therefore, depending solely on this heuristic may lead to suboptimal outcomes.

The initial comparison involved applying a threshold method based on the variance within the time-series. This method computes the variance across a sliding window along the signal. When the variance within the sliding window exceeds a predetermined threshold value, it indicates a transition from background noise to actual signal, thereby defining the onset time of the signal.

The choice to utilize the variance threshold instead of the conventional mean is driven by the characteristic of the analyzed signals, which possess a zero mean. Therefore, the variance-based method is deemed more effective in this context.

In this method, the parameters are represented by the size of the sliding window and the target value of the threshold. The definition of these parameters may represent a challenging task. To ensure an unbiased comparison with the method presented in this work, tests were conducted using various parameter values. Table 4 presents the results with the lowest error, obtained through experimentation with different parameter settings. Specifically, the table displays the outcomes for three sliding window sizes and a variance threshold of 6.5×10^{-3}.

In Table 4, the comparison is made based on the average Δt (10^{-6} s) between predicted and actual onset times.

Even after tuning the parameters of the threshold method, as depicted in Table 4, the results achieved with the U-Net approach are superior to those obtained using the threshold method across all analyzed parameter combinations.

In Table 4, when considering the optimal scenario with a sliding window size of 5, the deep learning method exhibits an average error that is less than half of that observed with the threshold method.

This can mainly be attributed to the noise present in the signal, which can lead to false detections in regions distant from the actual onset time. Furthermore, if the sliding window is too large, it may identify onset times prematurely, while if it is too small, there may not be sufficient data for a meaningful computation of the variance.

Further validation of the proposed methodology in this study was conducted by comparing the results with those obtained using the Change Point Method (CPM) (Ross et al., 2012; Ross, 2015) employing the Lepage test.

In this case, it was observed that the low variance inherent in acoustic signals made the Change Point Method (CPM) overly sensitive, leading to erroneous identification of noise as onset time.

In common practice, to reduce the method oversensitivity, a small amount of noise can be introduced to the data. In this case, 5% of the maximum fluctuation was used as background noise.

Nevertheless, even with this preprocessing step, the method exhibited the poorest performance, likely due to the intrinsic noise present in the signals used.

The results are presented in Table 4. A comparison based on the average Δt reveals that the error obtained with this well-established methodology is four times higher than that of the presented methodology.

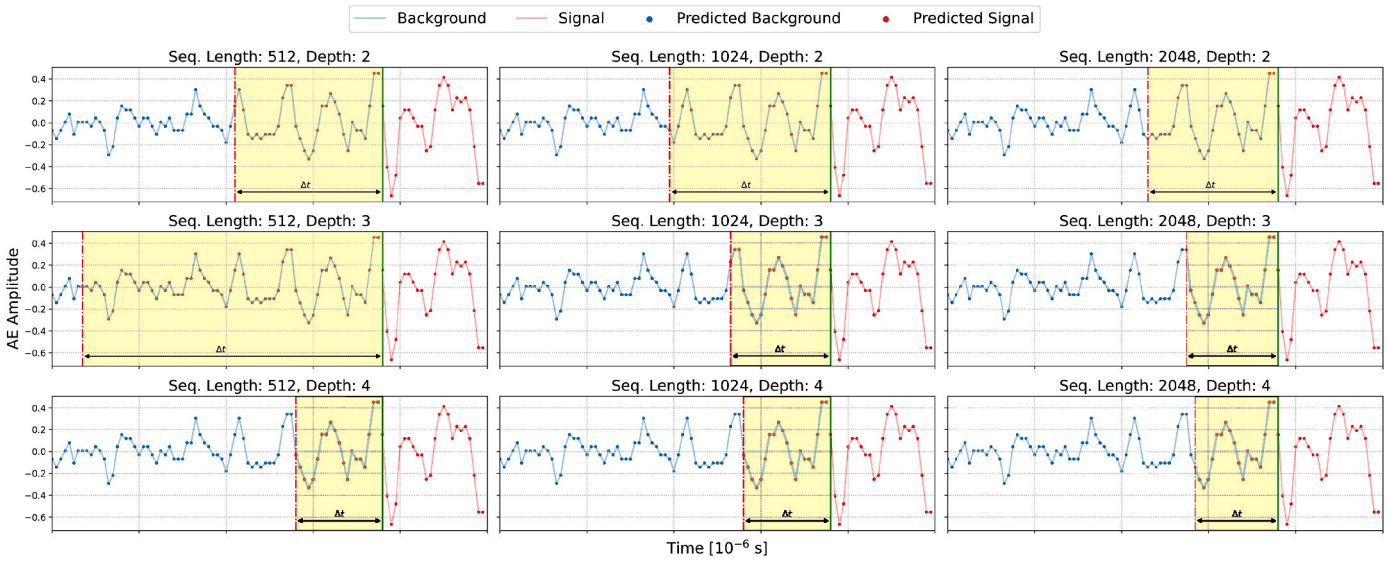

Fig. 6. Comparison of different U-Net models considering the distance between the predicted onset time and the ground truth.

Table 4
Comparison of different onset time detection methods in terms of average Δt (10^{-6} s) between predicted and actual onset times.

Method	$\Delta t[\mu s]$
Variance Thresholding (window size = 3)	25.46
Variance Thresholding (window size = 5)	22.50
Variance Thresholding (window size = 7)	30.47
CPM w/Lepage test	40.57
U-Net	10.62

Finally, Fig. 7 provides an example of the onset time identification on a signal from the dataset using the three compared methodologies. This visual comparison further enhances the superior accuracy of the method discussed in this paper.

3.2. Model validation on real AE signals

In this Section, the efficacy of the deep learning approach outlined in this study is demonstrated by evaluating it with a dataset comprising real acoustic emission signals. These signals were recorded through laboratory experiments conducted on concrete samples. Specifically, a three-point bending test was performed on a fiber-reinforced concrete specimen.

The tested sample, measuring 120 × 30 × 15 cm, included a centrally located notch with a depth of 5 cm. Loading was applied until failure, with control over the crack mouth opening displacement (CMOD). To record the acoustic emission waveforms emitted during fracture formation, a piezoelectric sensor was positioned on the specimen short side. Signals were sampled at a frequency of 1000 kHz.

The selection of the recorded time window is determined using a threshold amplitude method, as outlined in (Carpinteri et al., 2012). Here, x_t represents the time series. The threshold amplitude is established by comparing the average amplitude of a translated set of ten data points with four times the average amplitude of the signals from the first to the last recorded data point k. The amplitude threshold is expressed as follows:

$$\frac{\left(\sum_{t=k+1}^{10}|x_t|\right)}{10} \geq 4\frac{\left(\sum_{t=1}^{k}|x_t|\right)}{k} \tag{10}$$

The signals were recorded for a duration of 1024 μs, matching the length of signals in the training dataset. Subsequently, the acquired acoustic emission signals were tested using the method based on the U-Net architecture.

Fig. 8 demonstrates the method effectiveness in distinguishing background noise from the actual acoustic emission signals. Despite the inherent differences between signals generated by the PLB test and those from actual crack formation, the U-Net model trained solely on PLB test signals shows proficiency in identifying acoustic emission signals. However, it is important to highlight that the method was tested on signal segments trimmed to match the length of those used for training.

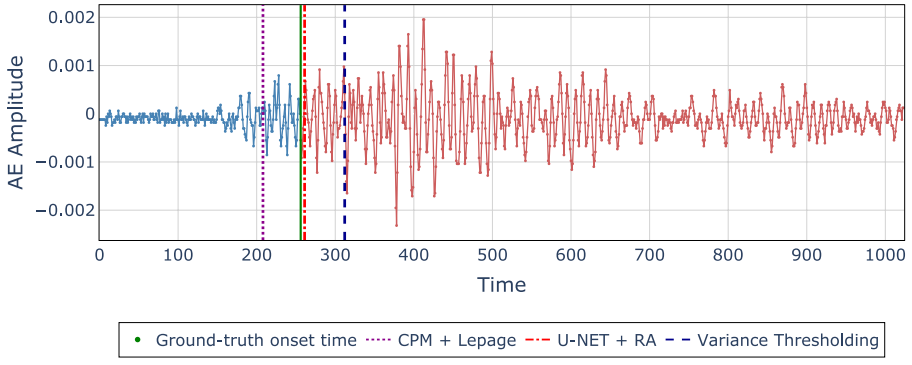

Fig. 7. Example of the onset time identification using the three compared methodologies. Evaluation of the U-Net model (depth of 4 and sequence length of 1024) against the variance thresholding (window size of 5 and threshold of 6.5 × 10^{-3}) and the CPM model with the Lepage test.

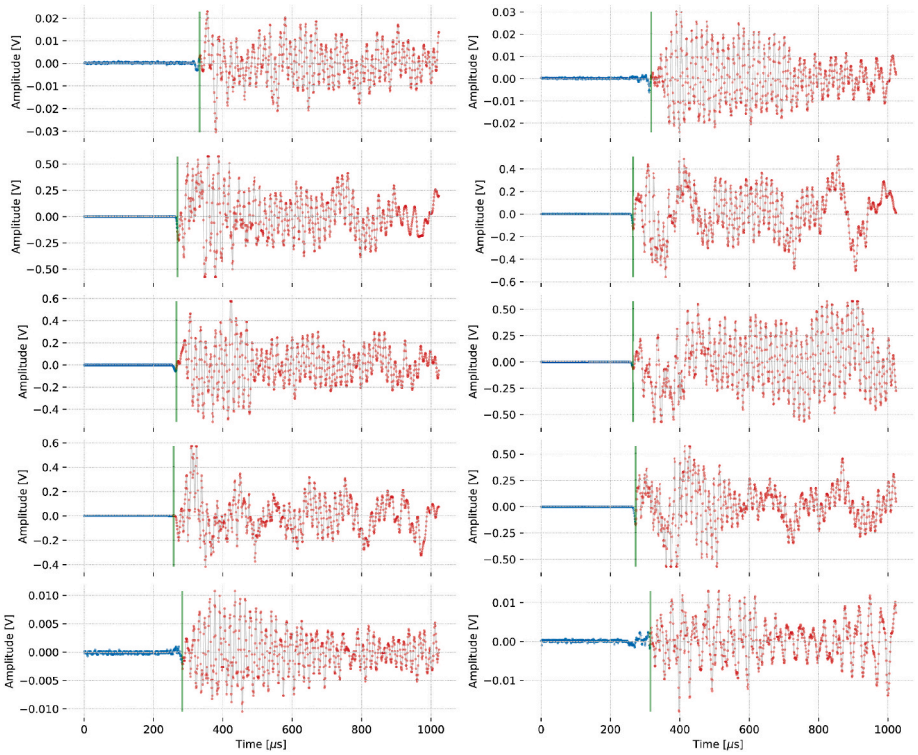

Fig. 8. Model validation on real AE signals. The onset time (green line) is identified when the classification label transitions from "background" (blue) to "signal" (red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

To further enhance accuracy and enable the use of continuous signals, model training using real acoustic emission data is advisable. Additionally, it is crucial to acknowledge that the tested acoustic emission signals were acquired in a laboratory setting, likely resulting in a lower background noise compared to field-recorded data.

Additionally, to assess the applicability of the presented methodology to continuous signals, a recording lasting 0.2 s was obtained during the three-point bending test, with an acquisition frequency of 1000 kHz. This recording captured two distinct acoustic emission signals, each characterized by significantly different amplitudes. The first signal peaked at approximately 0.6V, while the second signal was merely around 0.01V, representing less than 2% of the maximum amplitude.

Using the trigger method defined in (Carpinteri et al., 2012) or a conventional threshold method, it would have been challenging to detect the lower amplitude signal due to its minimal amplitude. However, as depicted in Fig. 9, the U-Net model successfully identified the signal.

In Fig. 10, a close-up view of the lower amplitude signal depicted in Fig. 9 is presented. The network successfully identifies the acoustic emission signal, despite the challenge posed by its significantly lower amplitude compared to the preceding signal.

In conclusion, the validation tests conducted on a small but real dataset have successfully confirmed the efficacy of the proposed methodology. The method demonstrated precise identification of onset times even on authentic acoustic emission signals. A natural progression for this methodology could involve direct training of the neural network using a diverse dataset of real signals. This would include signals with varying signal-to-noise ratios, reflecting operational monitoring scenarios in real-world field conditions. Such advancements could further enhance the robustness and applicability of the method.

4. Conclusions

This study presents a novel deep learning methodology for real-time

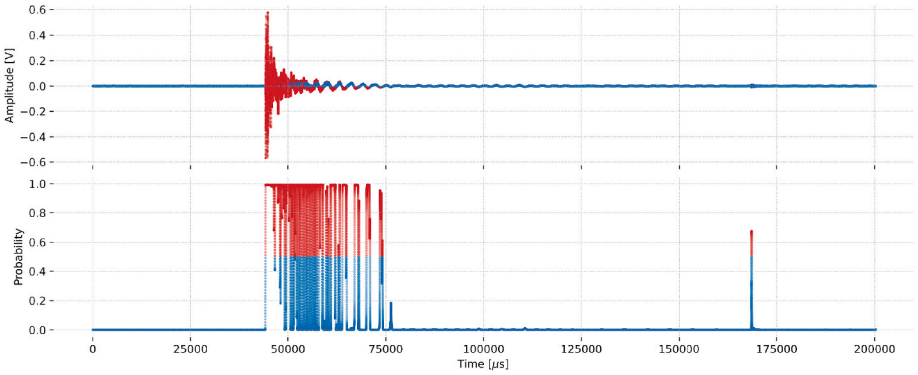

Fig. 9. Model validation on a real continuous AE signal lasting 0.2 s, demonstrating the effectiveness of the proposed method in detecting acoustic emissions of varying amplitudes. Classification and related probability of belonging to the "background" (blue) or "signal" (red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

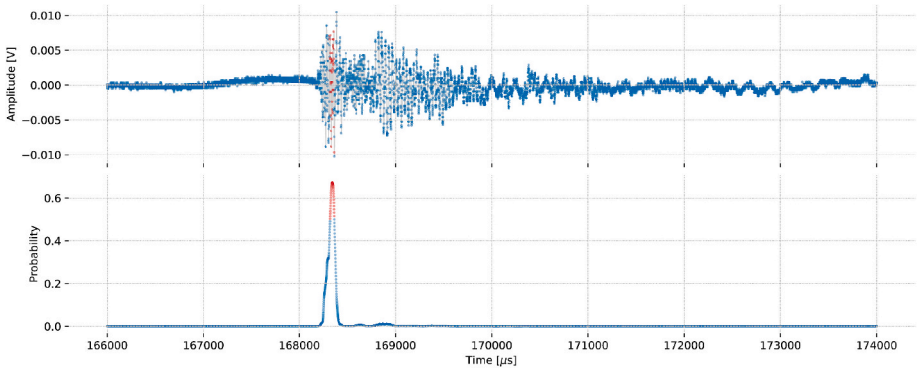

Fig. 10. Zoomed-in view of the lower amplitude signal from Fig. 9, demonstrating the ability of the proposed to detect subtle acoustic emission signals that may be difficult to identify with traditional methods.

identification of onset time in acoustic emission (AE) signals, with implications for structural health monitoring (SHM) in civil infrastructure. Through comprehensive experimentation on both simulated and real data, the efficacy of the proposed approach in accurately classifying signal samples and distinguishing between background noise and AE signals has been demonstrated.

The results of the experiments showcase the remarkable performance of the developed models, achieving high levels of accuracy across different evaluation metrics. Notably, the models exhibit minimal mean absolute error (MAE), indicating robust performance in accurately identifying onset times in AE signals. Moreover, the incorporation of a rolling average correction technique significantly enhances the precision of the proposed method, effectively reducing the average error between predicted and real onset times. Additionally, the method was compared with traditional methodologies for event detection widely used in signal processing, further validating its effectiveness.

Potential future advancements for this approach may involve training the neural network directly using a diverse dataset of real AE signals, including those with varying signal-to-noise ratios. This would allow for the simulation of conditions found in real-world operational monitoring scenarios. These developments hold promise for further enhancing the method resilience and practical applicability.

In summary, the findings of this study offer valuable insights into the application of deep learning techniques for AE signal analysis, advancing the field of SHM and paving the way for innovative solutions in structural defect detection and maintenance management. Further research in this area holds promise for continued improvements in accuracy and efficiency, ultimately benefiting the safety and resilience of built environments.

CRediT authorship contribution statement

Jonathan Melchiorre: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Data curation, Conceptualization, Project administration, Resources. **Leo D'Amato:** Writing – review & editing, Visualization, Validation, Software, Investigation, Formal analysis, Methodology. **Federico Agostini:** Visualization, Software, Methodology, Investigation, Formal analysis. **Antonino Maria Rizzo:** Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Aggelis, D.G., 2011. Classification of cracking mode in concrete by acoustic emission parameters. Mech. Res. Commun. 38 (3), 153–157.
Akaike, H., 1974. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. Ann. Inst. Stat. Math. 26, 363–387.
Anitescu, C., Atroshchenko, E., Alajlan, N., Rabczuk, T., 2019. Artificial neural network methods for the solution of second order boundary value problems. Comput. Mater. Continua (CMC) 59 (1), 345–359. https://doi.org/10.32604/cmc.2019.06641.
Arnau, A., et al., 2004. Piezoelectric Transducers and Applications, vol. 2004. Springer.
Baer, M., Kradolfer, U., 1987. An automatic phase picker for local and teleseismic events. Bull. Seismol. Soc. Am. 77 (4), 1437–1445.
Bai, F., Gagar, D., Foote, P., Zhao, Y., 2017. Comparison of alternatives to amplitude thresholding for onset detection of acoustic emission signals. Mech. Syst. Signal Process. 84, 717–730.
Bengio, Y., Goodfellow, I., Courville, A., 2017. Deep Learning, vol. 1. MIT press Cambridge, MA, USA.
Bilgen, M., Insana, M.F., 1998. Covariance analysis of time delay estimates for strained signals. IEEE Trans. Signal Process. 46 (10), 2589–2600.
Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning, vol. 4. Springer.
Boschetti, F., Dentith, M.D., List, R.D., 1996. A fractal-based algorithm for detecting first arrivals on seismic traces. Geophysics 61 (4), 1095–1102.
Buscema, M., 1998. Back propagation neural networks. Subst. Use Misuse 33 (2), 233–270.
Carpinteri, A., Lacidogna, G., Niccolini, G., 2006. Critical behaviour in concrete structures and damage localization by acoustic emission. In: Key Engineering Materials, vol. 312. Trans Tech Publ, pp. 305–310.
Carpinteri, A., Lacidogna, G., Manuello, A., 2007. Damage mechanisms interpreted by acoustic emission signal analysis. In: Key Engineering Materials, vol. 347. Trans Tech Publ, pp. 577–582.
Carpinteri, A., Xu, J., Lacidogna, G., Manuello, A., 2012. Reliable onset time determination and source location of acoustic emissions in concrete structures. Cement Concr. Compos. 34 (4), 529–537.
Carpinteri, A., Lacidogna, G., Invernizzi, S., Manuello Bertetto, A.D.B., et al., 2013. Ae monitoring and structural modeling of the asinelli tower in bologna. In: Proceedings of the 13th International Conference on Fracture (ICF13), China Science Literature. Publishing House, p. 234.
Chen, P.-H., Ding, J.-J., Huang, J.-Y., Tseng, T.-Y., 2020. Accurate onset detection algorithm using feature-layer-based deep learning architecture. In: 2020 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, pp. 1–5.
Ciaburro, G., 2020. Sound event detection in underground parking garage using convolutional neural network. Big Data and Cognitive Computing 4 (3), 20.
Ciaburro, G., Iannace, G., 2021. Modeling acoustic metamaterials based on reused buttons using data fitting with neural network. J. Acoust. Soc. Am. 150 (1), 51–63.
Ciampa, F., Meo, M., 2010. Acoustic emission source localization and velocity determination of the fundamental mode a0 using wavelet analysis and a Newton-based optimization technique. Smart Mater. Struct. 19 (4), 045027.
Du, G., Cao, X., Liang, J., Chen, X., Zhan, Y., 2020. Medical image segmentation based on u-net: a review. J. Imag. Sci. Technol.
Eaton, M.J., Pullin, R., Holford, K.M., 2012. Towards improved damage location using acoustic emission. Proc. IME C J. Mech. Eng. Sci. 226 (9), 2141–2153.
Emamian, V., Kaveh, M., Tewfik, A.H., Shi, Z., Jacobs, L.J., Jarzynski, J., 2003. Robust clustering of acoustic emission signals using neural networks and signal subspace projections. EURASIP J. Appl. Signal Process. 2003, 1–11.
Farrar, C.R., Worden, K., 2007. An introduction to structural health monitoring. Phil. Trans. Math. Phys. Eng. Sci. 365 (1851), 303–315.

Gorman, M.R., 1991. Plate wave acoustic emission. J. Acoust. Soc. Am. 90 (1), 358–364.
Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., et al., 2018. Recent advances in convolutional neural networks. Pattern Recogn. 77, 354–377.
Hinkley, D.V., 1971. Inference about the change-point from cumulative sum tests. Biometrika 58 (3), 509–523.
Ho, Y., Wookey, S., 2019. The real-world-weight cross-entropy loss function: modeling the costs of mislabeling. IEEE Access 8, 4806–4813.
Hua, Y., Guo, J., Zhao, H., 2015. Deep belief networks and deep learning. In: Proceedings of 2015 International Conference on Intelligent Computing and Internet of Things. IEEE, pp. 1–4.
Ince, N.F., Kao, C.-S., Kaveh, M., Tewfik, A., Labuz, J.F., 2010. A machine learning approach for locating acoustic emission. EURASIP J. Appl. Signal Process. 2010, 1–14.
Jierula, A., Oh, T.-M., Wang, S., Lee, J.-H., Kim, H., Lee, J.-W., 2021. Detection of damage locations and damage steps in pile foundations using acoustic emissions with deep learning technology. Front. Struct. Civ. Eng. 15, 318–332.
Kohonen, T., 1990. The self-organizing map. Proc. IEEE 78 (9), 1464–1480.
Kurz, J.H., Grosse, C.U., Reinhardt, H.-W., 2005. Strategies for reliable automatic onset time picking of acoustic emissions and of ultrasound signals in concrete. Ultrasonics 43 (7), 538–546.
Lacidogna, G., Manuello, A., Niccolini, G., Carpinteri, A., 2015. Acoustic emission monitoring of Italian historical buildings and the case study of the athena temple in syracuse. Architect. Sci. Rev. 58 (4), 290–299.
Lee, S., Ha, J., Zokhirova, M., Moon, H., Lee, J., 2018. Background information of deep learning for structural engineering. Arch. Comput. Methods Eng. 25, 121–129.
Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440.
Madarshahian, R., Soltangharaei, V., Anay, R., Caicedo, J.M., Ziehl, P., 2019a. Hsu-nielsen source acoustic emission data on a concrete block. Data Brief 23, 103813. https://doi.org/10.1016/j.dib.2019.103813. https://www.sciencedirect.com/science/article/pii/S2352340919301647.
Madarshahian, R., Ziehl, P., Caicedo, J.M., 2019b. Acoustic emission bayesian source location: onset time challenge. Mech. Syst. Signal Process. 123, 483–495.
Manuello Bertetto, A., Masera, D., Carpinteri, A., 2020. Acoustic emission monitoring of the turin cathedral bell tower: foreshock and aftershock discrimination. Appl. Sci. 10 (11), 3931.
Manuello Bertetto, A., Marmo, F., Melchiorre, J., 2023. Acoustic emission monitoring and thrust network analysis of the central nave vaults of the turin cathedral. In: Italian Workshop on Shell and Spatial Structures. Springer, pp. 241–249.
Manuello, A., Niccolini, G., Carpinteri, A., 2019a. Ae monitoring of a concrete arch road tunnel: damage evolution and localization. Eng. Fract. Mech. 210, 279–287.
Manuello, A., Masera, D., Carpinteri, A., 2019b. Ae damage assessment in the bell tower of the turin cathedral. Key Eng. Mater. 817, 579–585.
Manuello, A., Marmo, F., Melchiorre, J., 2024. Investigating and monitoring central nave vaults of the turin cathedral with acoustic emissions and thrust network analysis. Develop. Built Environ., 100434
Maradei, C., Piotrkowski, R., Serrano, E., Ruzzante, J., 2003. Acoustic emission signal analysis in machining processes using wavelet packets. Lat. Am. Appl. Res. 33 (4), 443–448.
Marasco, G., Rosso, M.M., Aiello, S., Aloisio, A., Cirrincione, G., Chiaia, B., Marano, G.C., 2022. Ground penetrating radar fourier pre-processing for deep learning tunnel defects' automated classification. In: International Conference on Engineering Applications of Neural Networks. Springer, pp. 165–176.
Melchiorre, J., Sardone, L., Rosso, M.M., Aloisio, A., 2022. Intelligent structural damage detection with mems-like sensors noisy data. In: International Conference on Communication and Intelligent Systems. Springer, pp. 631–642.
Melchiorre, J., Rosso, M.M., Cirrincione, G., Marano, G.C., 2023a. Compact convolutional transformer fourier analysis for gpr tunnels assessment. In: 2023 International Conference on Control, Automation and Diagnosis (ICCAD). IEEE, pp. 1–6.
Melchiorre, J., Rosso, M.M., Cucuzza, R., D'Alto, E., Manuello, A., Marano, G.C., 2023b. Deep acoustic emission detection trained on seismic signals. In: Applications of Artificial Intelligence and Neural Systems to Data Science. Springer, pp. 83–92.
Melchiorre, J., Manuello Bertetto, A., Rosso, M.M., Marano, G.C., 2023c. Acoustic emission and artificial intelligence procedure for crack source localization. Sensors 23 (2), 693.
Mesaros, A., Heittola, T., Virtanen, T., Plumbley, M.D., 2021. Sound event detection: a tutorial. IEEE Signal Process. Mag. 38 (5), 67–83.
I.N.M. MISTRAS Group, 1978. Mistras Group. https://mistrasgroup.com/. World Headquarters in Princeton Junction, NJ – USA.
Mitchell, T.M., 1997. Mach. Learn.
Niccolini, G., Carpinteri, A., Lacidogna, G., Manuello, A., 2011. Acoustic emission monitoring of the syracuse athena temple: scale invariance in the timing of ruptures. Phys. Rev. Lett. 106 (10), 108503.
Ohtsu, M., 1987. Acoustic emission characteristics in concrete and diagnostic applications. J. Acoust. Emiss. 6 (2), 99–108.
Ohtsu, M., Okamoto, T., Yuyama, S., 1998. Moment tensor analysis of acoustic emission for cracking mechanisms in concrete. Struct. J. 95 (2), 87–95.
Pasca, D.P., Aloisio, A., Rosso, M.M., Sotiropoulos, S., 2022. Pyoma and pyoma_gui: a python module and software for operational modal analysis. SoftwareX 20, 101216.
Patrick, M.K., Adekoya, A.F., Mighty, A.A., Edward, B.Y., 2022. Capsule networks–a survey. J. King Saud Univer.-comput. Inform. Sci. 34 (1), 1295–1310.
Priyanka, D. Kumar, 2020. Decision tree classifier: a detailed survey. Int. J. Inf. Decis. Sci. 12 (3), 246–269.
Rocchi, A., Santecchia, E., Ciciulla, F., Mengucci, P., Barucca, G., 2019. Characterization and optimization of level measurement by an ultrasonic sensor system. IEEE Sensor. J. 19 (8), 3077–3084.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.
Ross, G., 2015. Parametric and nonparametric sequential change detection in r: the cpm package. J. Stat. Software 66 (8). https://doi.org/10.18637/jss.v066.i03.
Ross, G., Tasoulis, D., Adams, N., 2012. Nonparametric monitoring of data streams for changes in location and scale. Technometrics 53, 379–389. https://doi.org/10.1198/TECH.2011.10069.
Rosso, M.M., Aloisio, A., Melchiorre, J., Huo, F., Marano, G.C., 2023a. Noise effects analysis on subspace-based damage detection with neural networks. Structures 54, 23–37. Elsevier.
Rosso, M.M., Aloisio, A., Parol, J., Marano, G.C., Quaranta, G., 2023b. Intelligent automatic operational modal analysis. Mech. Syst. Signal Process. 201, 110669 https://doi.org/10.1016/j.ymssp.2023.110669.
Ruby, U., Yendapalli, V., 2020. Binary cross entropy with deep learning technique for image classification. Int. J. Adv. Trends Comput. Sci. Eng. 9 (10).
Sadowsky, J., 1996. Investigation of signal characteristics using the continuous wavelet transform, johns hopkins apl technical digest, 17 (3), 258–269.
Scruby, C.B., 1987. An introduction to acoustic emission. J. Phys. E Sci. Instrum. 20 (8), 946.
Sedlak, P., Hirose, Y., Enoki, M., 2013. Acoustic emission localization in thin multi-layer plates using first-arrival determination. Mech. Syst. Signal Process. 36 (2), 636–649.
Siddique, N., Paheding, S., Elkin, C.P., Devabhaktuni, V., 2021. U-net and its variants for medical image segmentation: a review of theory and applications. IEEE Access 9, 82031–82057.
Sodsri, C., 2003. Time-varying Autoregressive Modelling for Nonstationary Acoustic Signal and its Frequency Analysis. The Pennsylvania State University.
Sohn, H., Farrar, C.R., Hemez, F.M., Shunk, D.D., Stinemates, D.W., Nadler, B.R., Czarnecki, J.J., 2003. A Review of Structural Health Monitoring Literature: 1996–2001, vol. 1. Los Alamos National Laboratory, USA, p. 16.
Sundararajan, D., 2001. The Discrete Fourier Transform: Theory, Algorithms and Applications. World Scientific.
Suthaharan, S., Suthaharan, S., 2016. Support Vector Machine, Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning, pp. 207–235.
Tapeh, A.T.G., Naser, M., 2023. Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices. Arch. Comput. Methods Eng. 30 (1), 115–159.
Thompson, J., Smith, B., Warner, A., Jot, J.-M., 2012. Direct-diffuse decomposition of multichannel signals using a system of pairwise correlations. In: Audio Engineering Society Convention, vol. 133. Audio Engineering Society.
Voulodimos, A., Doulamis, N., Bebis, G., Stathaki, T., 2018a. Recent developments in deep learning for engineering applications. Comput. Intell. Neurosci. 2018.
Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., et al., 2018b. Deep learning for computer vision: a brief review. Comput. Intell. Neurosci. 2018.
Wang, W., Siau, K., 2019. Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: a review and research agenda. J. Database Manag. 30 (1), 61–79.
Wang, J., Lu, S., Wang, S.-H., Zhang, Y.-D., 2022. A review on extreme learning machine. Multimed. Tool. Appl. 81 (29), 41611–41660.
Weinberg, A.I., Last, M., 2019. Selecting a representative decision tree from an ensemble of decision-tree models for fast big data classification. J. Big Data 6 (1), 1–17.
Weiss, A., Weinstein, E., 1983. Fundamental limitations in passive time delay estimation–part i: narrow-band systems. IEEE Trans. Acoust. Speech Signal Process. 31 (2), 472–486.
Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Syst. 2 (1–3), 37–52.
Zhang, M., Li, M., Zhang, J., Liu, L., Li, H., 2020. Onset detection of ultrasonic signals for the testing of concrete foundation piles by coupled continuous wavelet transform and machine learning algorithms. Adv. Eng. Inf. 43, 101034.
Ziola, S.M., Gorman, M.R., 1991. Source location in thin plates using cross-correlation. J. Acoust. Soc. Am. 90 (5), 2551–2556.
Zonzini, F., Bogomolov, D., Dhamija, T., Testoni, N., De Marchi, L., Marzani, A., 2022. Deep learning approaches for robust time of arrival estimation in acoustic emission monitoring. Sensors 22 (3), 1091.