Advancing deep learning-based acoustic leak detection methods towards application for water distribution systems from a data-centric perspective

Wu, Yipeng; Ma, Xingke; Guo, Guancheng; Jia, Tianlong; Huang, Yujun; Liu, Shuming; Fan, Jingjing; Wu, Xue

DOI
10.1016/j.watres.2024.121999

Publication date
2024

Document Version
Final published version

Published in
Water Research

Water Research 261 (2024) 121999

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Advancing deep learning-based acoustic leak detection methods towards application for water distribution systems from a data-centric perspective

Yipeng Wu [a], Xingke Ma [a], Guancheng Guo [a], Tianlong Jia [b], Yujun Huang [a], Shuming Liu [a,*], Jingjing Fan [c], Xue Wu [a]

[a] School of Environment, Tsinghua University, 100084, Beijing, China
[b] Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2628 CN Delft, the Netherlands
[c] Shanghai Lingang Water & Wastewater Development Co., Ltd., 201306, Shanghai, China

ARTICLE INFO

Key words:
Water distribution system
Acoustic leak detection
Data augmentation
Convolutional neural network
Time–frequency spectrogram

ABSTRACT

Against the backdrop of severe leakage issue in water distribution systems (WDSs), numerous researchers have focused on the development of deep learning-based acoustic leak detection technologies. However, these studies often prioritize model development while neglecting the importance of data. This research explores the impact of data augmentation techniques on enhancing deep learning-based acoustic leak detection methods. Five random transformation-based methods—jittering, scaling, warping, iterated amplitude adjusted Fourier transform (IAAFT), and masking—are proposed. Jittering, scaling, warping, and IAAFT directly process original signals, while masking operating on time-frequency spectrograms. Acoustic signals from a real-world WDS are augmented, and the efficacy is validated using convolutional neural network classifiers to identify the spectrograms of acoustic signals. Results indicate the importance of implementing data augmentation before data splitting to prevent data leakage and overly optimistic outcomes. Among the techniques, IAAFT stands out, significantly increasing data volume and diversity, improving recognition accuracy by over 7%. Masking enhances performance mainly by compelling the classifier to learn global features of the spectrograms. Sequential application of IAAFT and masking further strengthens leak detection performance. Furthermore, when applying a complex model to acoustic leakage detection through transfer learning, data augmentation can also enhance the effectiveness of transfer learning. These findings advance artificial intelligence-driven acoustic leak detection technology from a data-centric perspective towards more mature applications.

1. Introduction

Under the backdrop of climate change and urbanization, the situation of global water scarcity is becoming more and more serious. Over the past two decades, more than 80 metropolitan cities have experienced extreme drought and water shortages around the world (Zhang et al., 2019). A water distribution system (WDS) is a large-scale urban infrastructure responsible for the transportation of water resources, and the effectiveness of its operation and maintenance is crucial for the sustainable utilization of water resources. However, factors such as aging and corrosion of pipelines, external damage, and improper management have led to severe leakage issues across the world (Bozkurt et al., 2022). Water losses typically account for 20–30% of the water supply, and in countries with poor WDSs, it can even reach up to 50% (El-Zahab and Zayed, 2019). Except for the waste of water resources, leakage can result in the contamination of water, compromising water quality and putting public health at risk (Fox et al., 2016). The energy used to transport water, which is wasted due to leakage, further exacerbates the environmental impact and contributes to the emission of greenhouse gasses (Smith et al., 2018). Under these conditions, it is urgent to develop leakage management measures.

Acoustic leak detection is an effective method that has been widely applied in water companies for a long time. Inspection workers can use acoustic devices such as listening sticks to identify leak sound waves in pipelines by perceiving the amplitude and frequency of acoustic signals (C. Zhang et al., 2022). With technological advancements, acoustic monitoring devices like hydrophones and accelerators are now used in leak detection for WDSs (Vanijjirattikhan et al., 2022). An increasing number of researchers are focusing on the application of machine learning techniques to analyze high-frequency data collected by such

* Corresponding author.
E-mail address: shumingliu@tsinghua.edu.cn (S. Liu).

https://doi.org/10.1016/j.watres.2024.121999
Received 20 February 2024; Received in revised form 20 June 2024; Accepted 24 June 2024
Available online 24 June 2024

devices, aiming to achieve automated leak detection.

Numerous scholars have conducted extensive research around the issue of machine learning-based acoustic leak detection, gradually forming a two-stage framework known as "acoustic feature extraction-acoustic signal classification" (Fan et al., 2022). Acoustic feature extraction aims to extract features from both the time and frequency domains using methods such as energy analysis and Fourier transforms to highlight differences between leak signals and ambient noises. Acoustic signal classification, relying on the extracted features, utilizes various machine learning models to automatically classify acoustic signals into categories such as leak or non-leak. For instance, some researchers have extracted features such as energy and kurtosis factor in the time and frequency domains, developing classifiers like k-nearest neighbors, artificial neural networks, and gradient boosting trees for leak detection (Li et al., 2018; Ravichandran et al., 2021; Ullah et al., 2023). In addition to utilizing time and frequency domain statistical features, other scholars have proposed linear predictive features, accurately extracting resonance peaks from acoustic signals and effectively improved the detection performance (Cody et al., 2020; Guo et al., 2020).

With the rapid advancement of deep learning, the stage of acoustic feature extraction (typically requiring domain knowledge) can be realized automatically through the stacking of neural network layers (Fu et al., 2022). This provides an end-to-end solution for acoustic leak detection. The most popular approach involves using two-dimensional images, generated after applying short-time Fourier transform (STFT) or wavelet transform to acoustic time series, as inputs to convolutional neural networks (CNNs) for developing classifiers (Guo et al., 2021; Shukla and Piratla, 2020). Literature suggests that deep learning-based acoustic leak detection methods have shown excellent performance, with recognition accuracy often reaching 95% (Ahmad et al., 2022; Kang et al., 2018; Shukla and Piratla, 2020; Yu et al., 2023) or even 99% (Guo et al., 2021; Zhang et al., 2023). Guo et al. (2021) and Yu et al. (2023) have indicated that CNN classifiers outperform traditional two-stage methods.

The application of deep learning represents an improvement in acoustic leak detection technology from a modeling perspective. In recent years, data-centric artificial intelligence (AI) has gained great attention, emphasizing the importance of high-quality and large-scale data for machine learning models (Singh, 2023; Zha et al., 2023). Overall, in previous research, there has not been a strong emphasis on the importance of data. To construct deep learning models, the common practice in existing research has been to perform framing on a small amount of acoustic data (i.e., segmenting a complete signal into several frames) and using this augmented dataset for model development (Wu et al., 2023). Each framed segment is treated as distinct data, despite being generated under nearly identical conditions. However, some studies have highlighted that acoustic data in water pipelines are influenced by various factors such as pipeline fittings and environmental noises. Simulated data generated in the laboratory and the limited data collected from WDSs may not comprehensively reflect the characteristics of leak signals and ambient noises (Sitaropoulos et al., 2023). Models trained on these data may carry the risk of ineffectiveness in practical applications (Bakhtawar and Zayed, 2021). Despite the impressive results reported in the literature, few of these deep learning-based approaches have been documented to be successfully implemented on a large scale. Furthermore, Fares et al. (2023) showed that the accuracy of a well-trained deep learning-based leak detection model reduced by 4% when newly collected data are tested.

Data augmentation, which can effectively increase data quantity and diversity in situations where data is limited, is an important aspect of data-centric AI (Zha et al., 2023). In the field of deep learning for image recognition, data augmentation techniques involving random transformations like adding noise, cropping, and rotation, simulate real-world scenarios to increase data volume and diversity, preserving labels and enhancing generalization (Shorten and Khoshgoftaar, 2019; Alqudah et al., 2023). In deep learning-based acoustic leak detection, where creating a complete database with sufficient and accurately labeled data is costly in terms of time and labor, leveraging data augmentation before a comprehensive database is available could be essential to enhance performance. However, the application of data augmentation to high-frequency time series data, such as acoustic signals, presents unique challenges. The dynamic nature of these datasets, coupled with the sensitivity to environmental disturbances and the diversity of the targets being monitored, necessitates a more nuanced approach. Not all transformations that are effective for images are suitable for time series data, as they may inadvertently alter the fundamental characteristics of the signals, leading to misclassification or loss of critical information (Iwana and Uchida, 2021; Um et al., 2017). Nevertheless, in research focused on high-frequency sensor data (e.g., wearables, inertial measurement units), some data augmentation techniques have been explored. Certain techniques directly apply random transformations in the time domain, such as adding white noise and cropping to the original time series (Rashid and Louis, 2019; Um et al., 2017). Others transform high-frequency signals into the frequency domain or time-frequency domain for random variations (Park et al., 2019; Steven Eyobu and Han, 2018).

Given the aforementioned conditions, it remains unclear whether deep learning-based leak detection models have been correctly developed using framed data and whether there exist other appropriate methods for augmenting acoustic signals. To address these research gaps, this paper aims to explore suitable data augmentation techniques and enhance the applicability of deep learning-based acoustic leak detection methods in WDSs from a data-centric perspective. In the following sections, various data augmentation techniques will be introduced first. Subsequently, the discussion will revolve around the pitfalls and practical outcomes of applying these techniques, leading to the final conclusions.

2. Methodology

2.1. Data description and pre-processing

Data were gathered from accelerometers installed in an operational WDS located in a southern city of China. The accelerometers automatically recorded acoustic signals daily from 2:00 a.m. to 4:00 a.m. at a sampling frequency of 8192 Hz and a sampling duration of 5 s. A total of 600 signals were selected for this study, evenly split between non-leak and leak categories, effectively mitigating the impact of data imbalance. On-site inspection workers confirmed and labeled these data. Non-leak data predominantly consist of ambient noise and also include various environmental noises such as sounds of pumping and water dripping. The distance between any leak point and a sensor was within 200 m (constrained by the accelerometer's detection range). Skilled workers estimated flow rates at leak points, predominantly below 5 m^3/h. It is noteworthy that each leak signal in this study corresponds to a different leak point, with no repeated sampling. This ensures that the signals used for model evaluation have not appeared in the training data.

Constrained by the current data situation, this study exclusively focused on signals from metal pipes and the majority of the signals originated from ductile iron pipes with a diameter not exceeding DN 400. Accordingly, the deep learning classifiers used in this study (detailed information can be found in Section 2.3) have a specific application scope, being suitable for metal networks with relatively low levels of external noise, as the data was collected from metal pipes during the early morning hours without additional de-noising. However, this does not imply that the newly-proposed data augmentation techniques (detailed information can be found in Section 2.2) are solely applicable to metal pipes; they could also be extended to non-metal pipelines. Nonetheless, it is advisable to develop new classifiers tailored specifically for such materials, as acoustic signals from metal and non-metal pipelines exhibit diverse features (Fares et al., 2023).

Data pre-processing encompasses filtering and centralization. A Butterworth bandpass filter (50~2000 Hz) was employed to exclude undesirable components. The removal of trends in the initial data was achieved through data centralization, ensuring that each data sample has a mean of zero (Wu et al., 2023).

2.2. Data augmentation techniques

Considering that the commonly used inputs to deep learning classifiers for acoustic leak detection are images in the time-frequency domain, data augmentation techniques for both the original signals in the time domain and the transformed images in the time-frequency domain are introduced. Fig. 1 shows all the data augmentation techniques used in this paper. These methods aim to expand the dataset while preserving the labels of the original data as much as possible.

2.2.1. Data augmentation in time domain

(1) **Framing**. Framing, also known as slicing and analogous to cropping in image data augmentation, involves randomly extracting consecutive frames/slices from the original time series. The acoustic signals utilized in this study have a duration of 5 s. Within such a short period of time, it is reasonable to assume that the morphology of the leakage point, pipeline attributes, and external environment do not undergo sudden changes. Consequently, the labels (i.e., leak or non-leak) of the sliced frames remain consistent with the original time series. In this study, all acoustic signals are sliced into 10 non-overlapping frames, each representing a time series with 0.5 seconds' worth of data (i.e., 4096 data points). As illustrated in Fig. 1, framing is considered a preceding or subsequent step in all other data augmentation techniques.

(2) **Jittering**. Jittering, a straightforward transformation-based data augmentation technique, involves adding white noise to the original time series. Typically, this process entails adding a random value conforming to a Gaussian distribution to each time step. However, the magnitudes of acoustic signals exhibit significant variations due to diverse leak volumes, pipeline attributes, and other factors. Hence, this study employs the signal-to-noise ratio (SNR) to determine the appropriate level of added white noise for different signals. The SNR of a signal with jittering is computed as:

$$\mathrm{SNR} = 10\log_{10}\left(\sum_{n=0}^{N-1} s(n)^2 \Big/ \sum_{n=0}^{N-1} d(n)^2\right) \tag{1}$$

where $s(n)$ is the original signal, $d(n)$ is the added white noise, and N is the length of $s(n)$ and $d(n)$. For a given signal and a predefined SNR, the corresponding white noise can be determined. In pipelines, both the internal and external environments are complex and Guo et al. (2021) used relatively low SNRs (no more than 10 dB) to simulate signals affected by uncertainties. In this study, the SNR value ϵ for each signal conforms to a Gaussian distribution $\epsilon \sim \mathrm{N}(20,\ 10^2)$, which is generally higher than, yet inclusive of, the SNRs reported in the literature. This approach aims to simulate realistic environmental interference to a certain extent while attempting to preserve the original signal label. Jittering is applied before framing, as illustrated in Fig. 1.

(3) **Scaling**. Scaling involves altering the overall magnitude of an acoustic signal by multiplying it with a random scalar value while maintaining the original labels (Rashid and Louis, 2019). Using the scaling parameter α, the entire time series undergoes a multiplication by α. The scaling parameter α can be determined by a Gaussian distribution $\alpha \sim \mathrm{N}(1,\ \sigma^2)$. It is important to note that the concept of "scaling" in the context of time series differs from that in the image domain. In time series, it specifically refers to increasing the magnitude of individual elements without expanding the time series itself. The value of σ is set to 0.2 according to the study by Um et al. (2017). Scaling is applied prior to framing, as depicted in Fig. 1.

(4) **Permutation**. Permutation, as a method of data augmentation, involves rearranging segments of a time series to create a new pattern. The resulting data, known as surrogate data, maintains the mean and variance of the original time series after undergoing a random shuffle. However, it is crucial to note that the power spectrum, the main basis for manual acoustic leak detection, may undergo significant changes in surrogate data, potentially leading to alterations in the labels of the original time series (Lee et al., 2019). To address this limitation, Lee et al. (2019) introduced the iterated amplitude-adjusted Fourier transform (IAAFT) as a technique for time series augmentation, achieving promising results in the classification of data related to human arm motion monitoring.

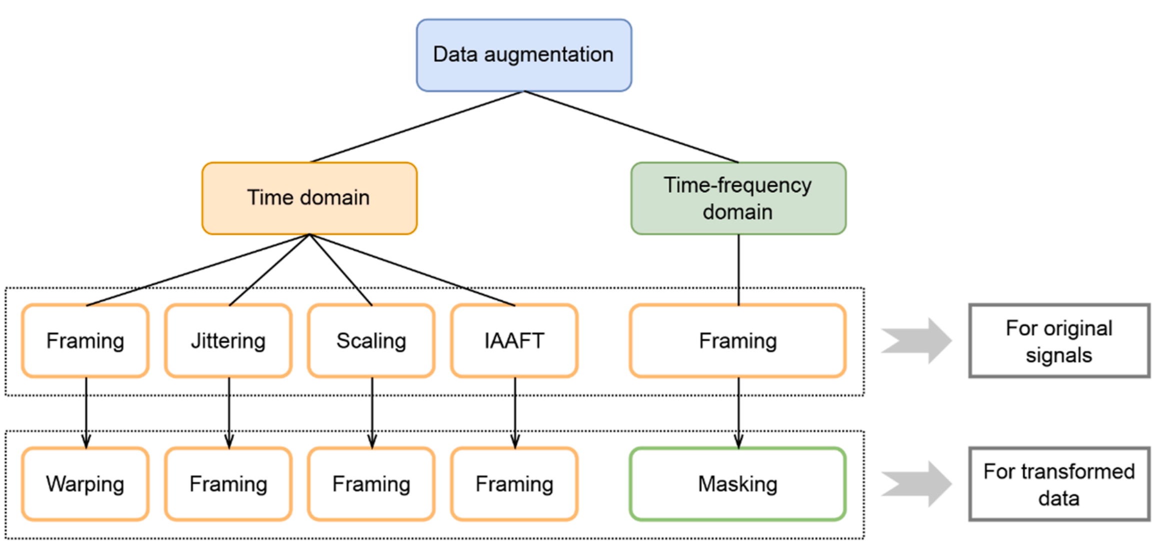

Fig. 1. Overview of data augmentation techniques. IAAFT= iterated amplitude-adjusted Fourier transform.

The IAAFT algorithm initiates with a random shuffle of data points, followed by iterative implementation of spectral adaptation and amplitude adaptation. During spectral adaptation, the Fourier transform of the initial or iterated time series is computed, and the magnitudes of its coefficients are replaced with those of the original time series. This process ensures that the new time series possesses the same power spectrum as the original. In the amplitude adaptation phase, adjustments are made based on amplitude ranking. For instance, the highest value in the iterated time series is substituted with the highest measured value in the original time series. As amplitude adjustment can influence the power spectrum, both amplitude and spectral adjustments are repeated until a convergence criterion is met. Further details and calculations regarding IAAFT can be found in the study by Venema et al. (2006). IAAFT is applied prior to framing, as depicted in Fig. 1.

(5) Warping. Warping involves perturbing a pattern within the original time series. One popular warping technique, known as window warping, selects a random time range and either compresses or extends it while keeping the remaining time range unchanged (Le Guennec et al., 2016). Following the original method's description, a chosen data segment within the time range is stretched by a factor of 2 or contracted by a factor of 0.5, thereby altering the length of the time series. To maintain the original length, this study divides a signal into 16 non-overlapping slices and randomly selects two slices to stretch by a factor of 1.5 and contract by a factor of 0.5, respectively. This approach perturbs the time series while preserving the majority (seven eighths) of it unchanged, thereby avoiding alterations to the label of the time series. It is important to note that warping is applied to sliced frames, as illustrated in Fig. 1.

The aforementioned methods, although not exclusively time-domain data augmentation techniques, are among the most commonly employed in the literature. Rotation is also often mentioned but has been demonstrated to be less effective for various types of time series, including high-frequency time series (Iwana and Uchida, 2021; Rashid and Louis, 2019). Consequently, this study omits the use of rotation.

2.2.2. Data augmentation in time-frequency domain

As shown in Fig. 1, each signal slice undergoes transformation into a spectrogram through STFT following framing. The spectrogram provides insights into the joint time-frequency distribution of acoustic signals, influenced by two crucial factors: frequency resolution F (equal to the number of rows in the spectrogram) and time resolution T (equal to the number of columns in the spectrogram). In the application of STFT, several key parameters must be determined, including the window function (e.g., rectangular or Hanning window), fast Fourier transform length (l, typically specified as a power of 2 and equal to frame length), and frame shift length (s, generally specified as a multiple of 1/4 the frame length). The resolution of a spectrogram φ is computed as:

$$\varphi = F \times T = \left(\frac{l}{2} + 1 - \delta\right) \times \left(\frac{N - l}{s} + 1\right) \quad (2)$$

where δ is a trimming factor to crop the signal spectrum to within 1∼2048 Hz. Considering the sampling frequency is 8192 Hz, $\delta = l/4$ in this study.

An augmentation method that acts on the spectrogram directly is proposed inspired by the work of Park et al. (2019). The objective of this method is to encourage a classifier to focus on more holistic features rather than just the localized aspects of the spectrograms, enhancing the robustness of the trained classifier against partial loss of frequency or temporal information within the time-frequency spectrogram. The approach involves augmenting the data by randomly masking specific information. To be specific, n_F rows and n_T columns are randomly chosen from each $F \times T$ spectrogram, and the values of the selected rows/columns are then set to zero.

In this study, a Hanning window was used to estimate the spectrum and fast Fourier transform length l and frame shift length s are 1024 and 256 respectively. Accordingly, the frequency resolution F and time resolution T of spectrograms are 257 and 13 respectively. The optimal values of n_F and n_T need to be determined through trial-and-error.

2.3. Deep learning classifiers for leak detection

CNN classifiers are the most widely employed deep learning models in the acoustic leak detection task. As illustrated in Fig. 2, a typical CNN classifier comprises convolution layers, pooling layers, fully connected layers, and an output layer. In a convolution layer, convolution kernels traverse a two-dimensional spectrogram, multiplying point-wise with the elements of each region, and the results are aggregated to obtain a value. These values, generated by the sliding convolution kernels, collectively form a feature map. Pooling layers serve as feature compressors for the generated feature maps after convolution. This compression aids in reducing computational complexity and extracting essential information from the convolution output features. Various combinations between convolution and pooling layers are possible, including one-to-one or many-to-one configurations. Fig. 2 illustrates the one-to-one combination. Fully connected layers consolidate feature maps to generate informative data for classification purposes. The output layer, equipped with a single neuron, applies a sigmoid function, with the output representing the probability of a leak occurrence.

In this study, the evaluation of data augmentation techniques is mainly based on a time–frequency CNN (TFCNN) classifier, which features two convolutional layers, one max-pooling layer, two fully connected layers (with 512 and 256 neurons, respectively), and an output layer. TFCNN is a relatively simple CNN model and its architecture is referenced from the acoustic leak detection study by Guo et al. (2021).

Increasing the depth of CNN models by adding more convolutional layers is a key way to enhance classifier performance (Tan and Le, 2019). However, deeper models require significantly more parameters, making them difficult to train from scratch with limited data. To address this, using transfer learning to apply complex models from other domains to a target task (e.g., acoustic leak detection) has become common. This approach allows for training on the target task with a small amount of new data, using pretrained parameters (P. Zhang et al., 2022). Nevertheless, data augmentation is also indispensable for ensuring model effectiveness in transfer learning (Chollet, 2023). Therefore, this study also investigates the impact of data augmentation on transfer learning in the acoustic leak detection task.

In this study, the model selected for transfer learning is the pretrained audio neural network (PANN), a CNN classifier with 10 convolutional layers developed by Kong et al. (2020). Originally trained on the large audio database AudioSet for the purpose of audio classification, the PANN model is well-suited for transfer to acoustic leak detection due to the task similarity and the use of spectrograms as input data. This model also includes the fundamental components depicted in Fig. 2. For a detailed description of its architecture, please refer to Kong et al. (2020).

2.4. Performance evaluation of leak detection

Accuracy, sensitivity, specificity, and receiver operating characteristic (ROC) curve and its area under curve (AUC) are commonly used to evaluate the performance of classifiers (Guo et al., 2020; Yu et al., 2023). In this paper, leak data are considered positive samples and non-leak data are considered negative samples. The corrected detected leak data are true positive (TP), the undetected leak data are false negative (FN), the wrongly detected non-leak data are false positive (FP), and the remaining ones are true negative (TN). Accuracy, sensitivity, specificity are computed as follows:

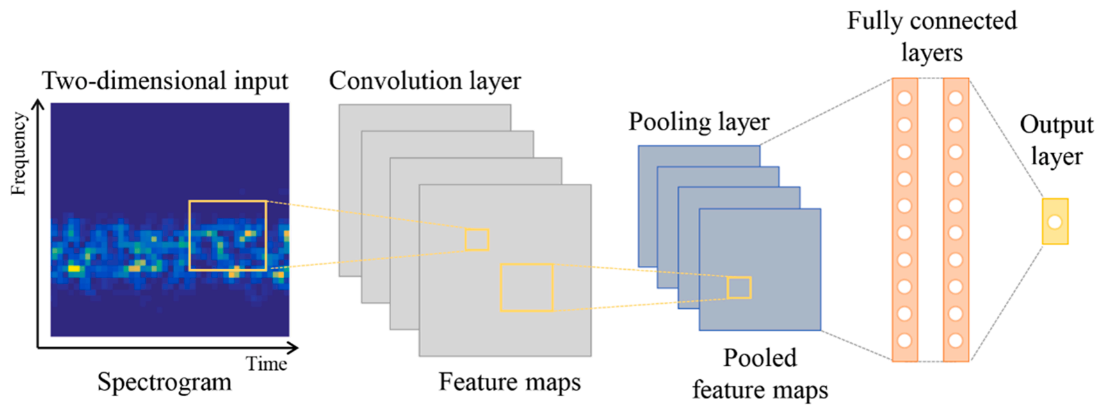

Fig. 2. Diagram of the main components of a CNN classifier.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{3}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \tag{5}$$

Accuracy describes the ratio of the number of correctly identified signals to the total number of signals. Sensitivity is the ability to correctly detecting real leak signals, while specificity is the ability to avoid wrongly detecting non-leak signals as leak ones. Obviously, the larger the accuracy, sensitivity, and specificity, the better the performance of a leak detection classifier. For a well-trained leak detection classifier, increasing sensitivity comes at the cost of decreasing specificity. The ROC curve describes the relationship between 1-specificity (x axis) and sensitivity (y axis) when the method adjusts its classification threshold. The larger the AUC value corresponding to the curve, the better a method. Accuracy, sensitivity, and specificity are computed based on a default classification threshold 0.5.

3. Results and discussion

3.1. Method implementation

3.1.1. Implementation of data augmentation

In this study, 300 leak signals and 300 non-leak signals were collected on-site. To establish a base classifier and facilitate comparisons, these data underwent augmentation using the framing technique, resulting in 3000, 1500, and 1500 frames for training, validation, and testing, respectively. Unlike the conventional data splitting ratios (e.g., 8:1:1), a different ratio of 2:1:1 was employed due to the use of data augmentation.

Section 2.2 details the generation of synthetic data through various augmentation techniques, aiming to increase the volume of signals while preserving labels. As framing is a precursor or subsequent step in other augmentation methods, five distinct techniques were applied: jittering, scaling, warping, IAAFT, and masking. These techniques are exclusively applied to the training signals, producing five types of datasets with varying data volumes—specifically, 6000, 9000, 12,000, 15,000, and 18,000 (i.e., with 3000, 6000, 9000, 12,000, and 15,000 augmented data) for each technique. Each type of dataset comprises 10 randomly generated subsets, resulting in a total of 50 datasets for each data augmentation technique. Consequently, classifiers will undergo ten times for each dataset size to mitigate the impact of randomness on a data augmentation technique's performance evaluation. Note that for the TFCNN, the optimal values of n_F and n_T in masking are 64 and 3, respectively, while for PANN, the optimal values are 16 and 1. All data augmentation methods were executed in a Python 3.7.13 environment and spectrograms were generated using the voicebox of MATLAB R2022b.

3.1.2. Implementation of deep learning classifiers

These deep learning classifiers, developed using Tensorflow 2.9.0, were implemented in the Python 3.7.13 environment. The activation function ReLU and the Adam optimizer were employed. To prevent overfitting, early-stopping and dropout mechanisms were incorporated. Detailed hyper-parameters are outlined in the subsequent sections. Hyper-parameter tuning was performed using Neural Network Intelligence (NNI), a Microsoft open-source toolkit operable in the Python environment (Microsoft, 2021). It is important to note that the hyper-parameters were exclusively tuned based on the dataset employing framing alone, to exclude performance improvements resulting from hyper-parameter optimization in the evaluation of data augmentation techniques. The deep learning classifiers ran on a server equipped with a GPU, with configurations including an Intel(R) Xeon(R) Gold 5218 CPU @ 2.30 GHz, 64 cores and 64 threads, 128 GB of RAM, and an NVIDIA GeForce RTX 3090Ti with 24 GB of graphics memory.

Prior to evaluating the data augmentation techniques, this study analyzed the sensitivity of the TFCNN and PANN models to varying data volumes. The results demonstrate that leak detection performance improves with increased data volume, underscoring the benefits of providing more data for deep learning models. Therefore, it is reasonable to assume that augmenting data to provide additional samples can enhance leak detection effectiveness. Detailed information on the sensitivity analysis can be found in the supplementary material.

3.2. Biased results caused by data leakage

In the current research on deep learning-based acoustic leak detection, framing is commonly employed. However, most studies did not explicitly describe whether framing was applied before splitting original signals into training, validation, and testing datasets or not. In other words, the impact of the order of implementation regarding data splitting and framing on detection performance has not been thoroughly discussed. This section aims to compare the performance of TFCNN classifiers under the two conditions mentioned above. The TFCNN classifier using data with framing implemented first is referred to as TFCNN-post-splitting, while the TFCNN classifier using data with data splitting implemented first is termed TFCNN-pre-splitting. The hyper-parameters of these classifiers are detailed in Table 1.

As illustrated in Fig. 3, the performance of TFCNN-post-splitting significantly outperforms that of TFCNN-pre-splitting, underscoring the critical importance of the implementation order in data splitting and framing. The essence of the implementation order hinges on whether each frame can be regarded as an entirely new sample. In the case of

Table 1
Optimal hyper-parameters of deep learning classifiers.

Model	Hyper-parameters
TFCNN-post-splitting	First convolutional layer: filters = 32, kernel size = (5, 5), strides = (4, 4) Second convolutional layer: filters = 64, kernel size = (3, 3), strides = (2, 2) Max-pooling layer: pooling size = (2, 2), strides = (2, 2) Learning rate = 0.008, batch size = 64, dropout rate = 0.1, epochs = 200
TFCNN-pre-splitting	First convolutional layer: filters = 32, kernel size = (5, 5), strides = (4, 4) Second convolutional layer: filters = 64, kernel size = (3, 3), strides = (2, 2) Max-pooling layer: pooling size = (2, 2), strides = (2, 2) Learning rate = 0.002, batch size = 256, dropout rate = 0.25, epochs = 200
PANN	Learning rate = 0.0005, batch size = 128, epochs = 200

Note: Model architecture and other hyper-parameters of PANN are consistent with the study by Kong et al. (2020).

TFCNN-post-splitting, each frame, whether derived from the same original signal or different signals, is treated as a distinct sample. Consequently, all frames can be randomly allocated to training, validation, and testing datasets. Conversely, for TFCNN-pre-splitting, frames originating from the same original signal are handled cautiously due to the likelihood of preserving highly similar characteristics. When frames from the same signal are concurrently distributed across training, validation, and testing datasets, the classifier may glean information from unavailable or unseen data during the training phase, resulting in biased optimistic outcomes. This phenomenon is commonly referred to as data leakage (Zhu et al., 2023).

To assess the occurrence of data leakage, a widely used similarity measure, cosine similarity (Wu et al., 2016), was employed to analyze the time-frequency spectrograms (flattened to vectors) of augmented frames. Fig. 4 illustrates the distribution of cosine similarity between spectrograms of different frames from the same signal (i.e., intra-similarity) and between frames from different signals (i.e., inter-similarity). This distribution plot was generated using kernel density estimation (Heidenreich et al., 2013). It is evident that, for both leak and non-leak signals, intra-similarity is centered around 0.5 and significantly surpasses inter-similarity, with a density peak close to 0. This observation indicates the occurrence of data leakage in the development of TFCNN-post-splitting, contributing to the surprisingly favorable results obtained. While performance comparisons between different classifiers based on post-splitting data processing may offer some guidance for model selection, the selected classifier is likely to perform poorly when detecting newly collected signals. Consequently, all data augmentation techniques are implemented after data splitting in the subsequent sections. It is also worth mentioning that the low

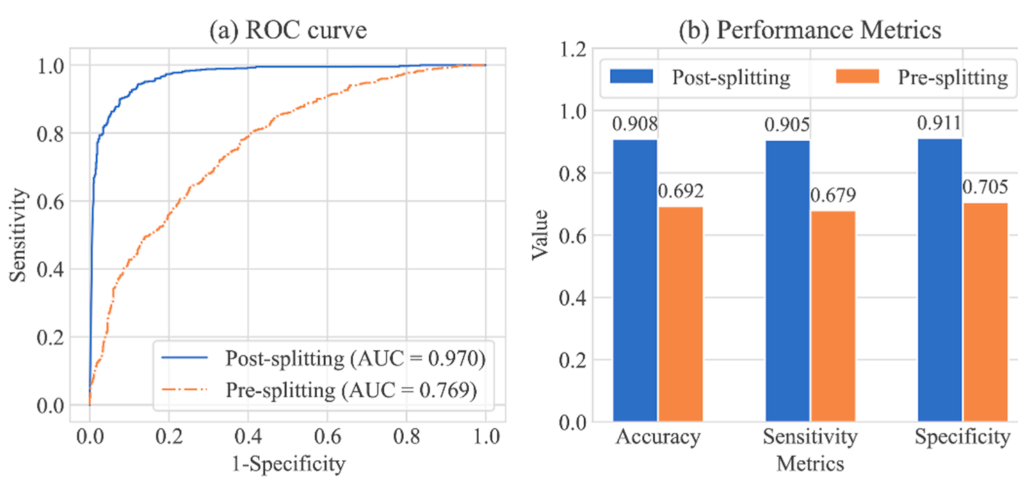

Fig. 3. Performance of TFCNN classifiers for two conditions where one implemented framing after splitting original signals into training, validation, and testing datasets (i.e., pre-splitting), while the other one implemented framing before splitting original signals into training, validation, and testing datasets (i.e., post-splitting).

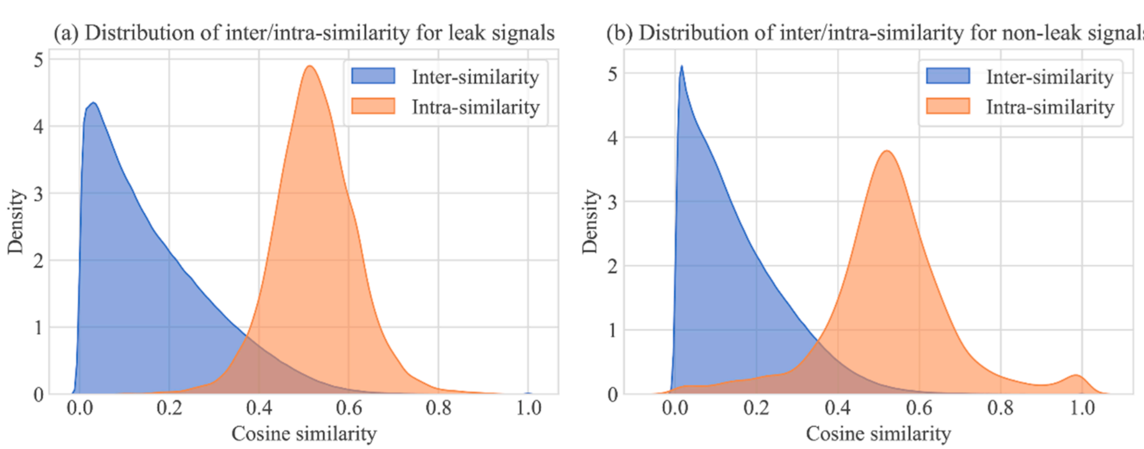

Fig. 4. Distribution of cosine similarity between spectrograms of different frames from the same signal (i.e., intra-similarity) and between frames from different signals (i.e., inter-similarity).

inter-similarity further proves that the signals used in this study come from different scenarios (e.g., different leak points).

3.3. Evaluation of individual data augmentation techniques

This section assesses the performance of leak detection using different individual data augmentation techniques. The baseline for

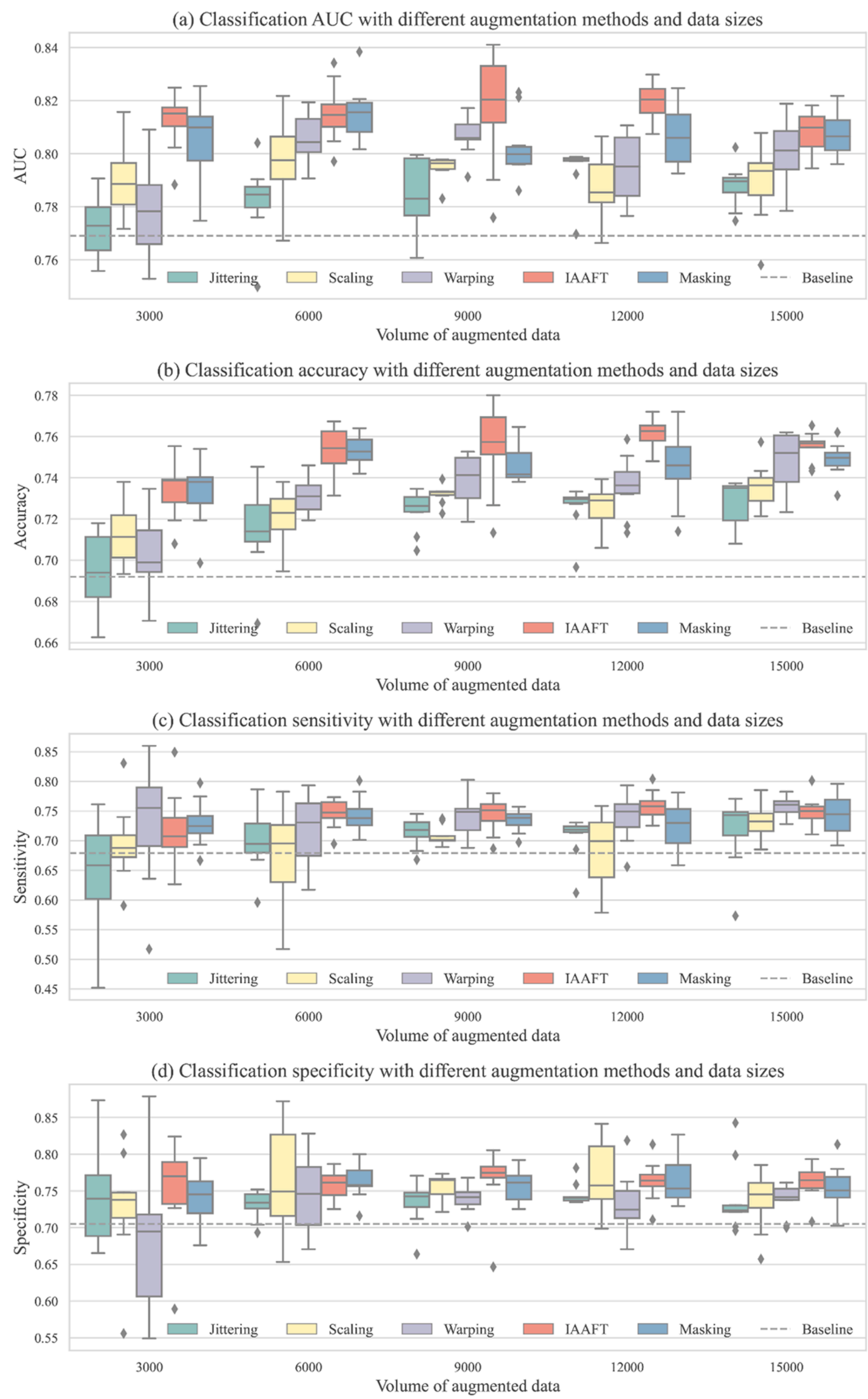

Fig. 5. Comparison of leak detection performance when using different single data augmentation techniques.

comparison is the performance of TFCNN-pre-splitting using framed data, with an AUC of 0.769, accuracy of 0.692, sensitivity of 0.679, and specificity of 0.705. As depicted in Fig. 5, the five data augmentation techniques demonstrate varying capabilities in terms of performance enhancement. With an increase in data volume, each technique contributes to an improvement in leak detection performance across all four performance metrics. The performance does not consistently improve with increasing data volume; in fact, there was a slight decrease after augmenting the data to 12,000 samples. This indicates that there is a limit to the effectiveness of data augmentation when the original dataset size is fixed, and more augmentation does not necessarily lead to better outcomes. Overall, IAAFT and masking stand out as the top-performing techniques. The accuracy (median value of ten runs) for IAAFT and masking increases by up to 0.071 and 0.061, respectively, when using 12,000 and 6000 augmented data. Detailed changes in the corresponding performance metrics for each data augmentation technique at different data volumes are presented in Table S1 in the supplementary material.

To explain why the five data augmentation methos show various abilities to enhance leak detection performance, the cosine similarity between the spectrograms of augmented data and original signals for different data augmentation techniques are calculated, which are illustrated in Fig. 6. For jittering and scaling, the cosine similarity between the spectrograms of augmented data and original signals is significantly high and close to 1, indicating that the augmented data have little difference with the original ones. Consequently, these two techniques bring limited additional information from the vast added data, leading to the least generalization ability gained by the deep learning classifier. This finding conforms to the work of Iwana and Uchida (2021). The other three techniques not only increase the amount of data but also bring about variability between the data. This helps alleviate overfitting and enhances the model's adaptability to diverse data.

It is straightforward to comprehend that data generated through masking exhibits low similarity with the original signals. This is due to the direct replacement of approximately 25% of information in both the time and frequency domains with zeros. Apart from introducing randomness to increase the diversity of the data, masking part of the spectrogram information enhances the robustness of classifiers against local variations in signals. By incorporating masking, the model is compelled to concentrate on the uncovered portions of the spectrogram, facilitating the learning of more global and holistic features. This aids classifiers in generalizing better to unseen data (Park et al., 2019). Consequently, masking visibly improves the performance of TFCNN-pre-splitting.

The augmented data generated by IAAFT also exhibits significant differences in the spectrogram compared to the original signal. This is attributed to IAAFT altering the arrangement order of the original signal, resulting in substantial changes in the time-frequency spectrogram, as depicted in Fig. 7. However, based on the principles of this technique, it is known that it ensures the time-domain statistical properties and spectrum of the augmented data closely match those of the original signal (Venema et al., 2006). This largely guarantees that the augmented data and the original data belong to the same category. Accordingly, IAAFT demonstrates the best performance. This result also indicates the importance of thorough analysis and understanding of the data characteristics through domain knowledge before designing data augmentation schemes.

While the augmented spectrograms produced by warping show relatively low similarity (with a median value of 0.52) to the original ones, the observed performance enhancement is not as pronounced as that achieved by IAAFT and masking. Additionally, warping lacks the robust evidence, as seen in IAAFT, of generating augmented data without altering data labels. Therefore, IAAFT and masking are the recommended techniques in this study.

3.4. Evaluation of combined data augmentation techniques

This section investigates whether the combined use of multiple data augmentation techniques can further enhance leak detection performance. The two most effective data augmentation methods, IAAFT and masking, are combined in two ways. The first approach involves generating synthetic data simultaneously using both techniques and merging the two types of augmented data to create a dataset. This is referred to as the parallel manner in this paper. The second approach, termed the sequential manner, first applies IAAFT and subsequently applies masking to the augmented data. Fifty datasets were generated for each of these combined techniques, and their ability to improve leak detection performance is illustrated in Fig. 8. Detailed changes in the corresponding performance metrics for each combined data augmentation technique at different data volumes are presented in Table S2 in the supplementary material.

The combined data augmentation technique in the sequential manner demonstrates a modest enhancement in leak detection performance, albeit not statistically significant. As illustrated in Fig. 6, the generated dataset exhibits the least similarity to the original signals when compared to any individual data augmentation technique. The accuracy (median value of ten runs) for this technique can increase by up to 0.075 when using both 9000 and 12,000 augmented data points

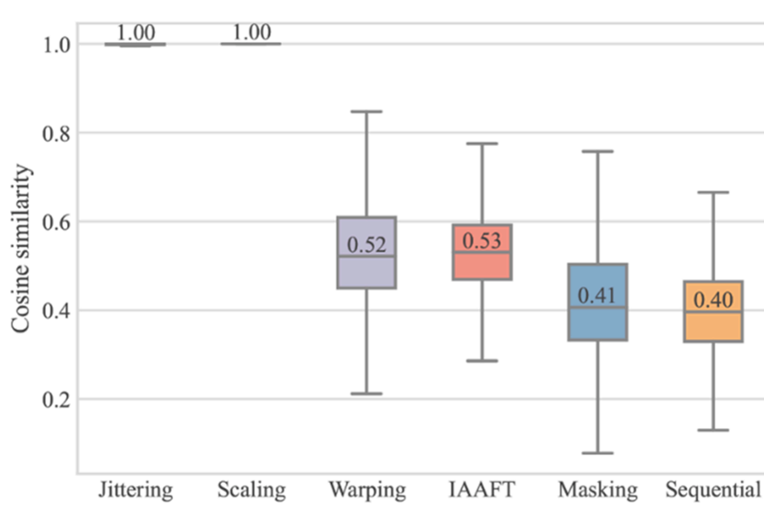

Fig. 6. Cosine similarity of spectrograms between augmented data and original signals for different data augmentation techniques. The numbers in the boxes represent the median values.

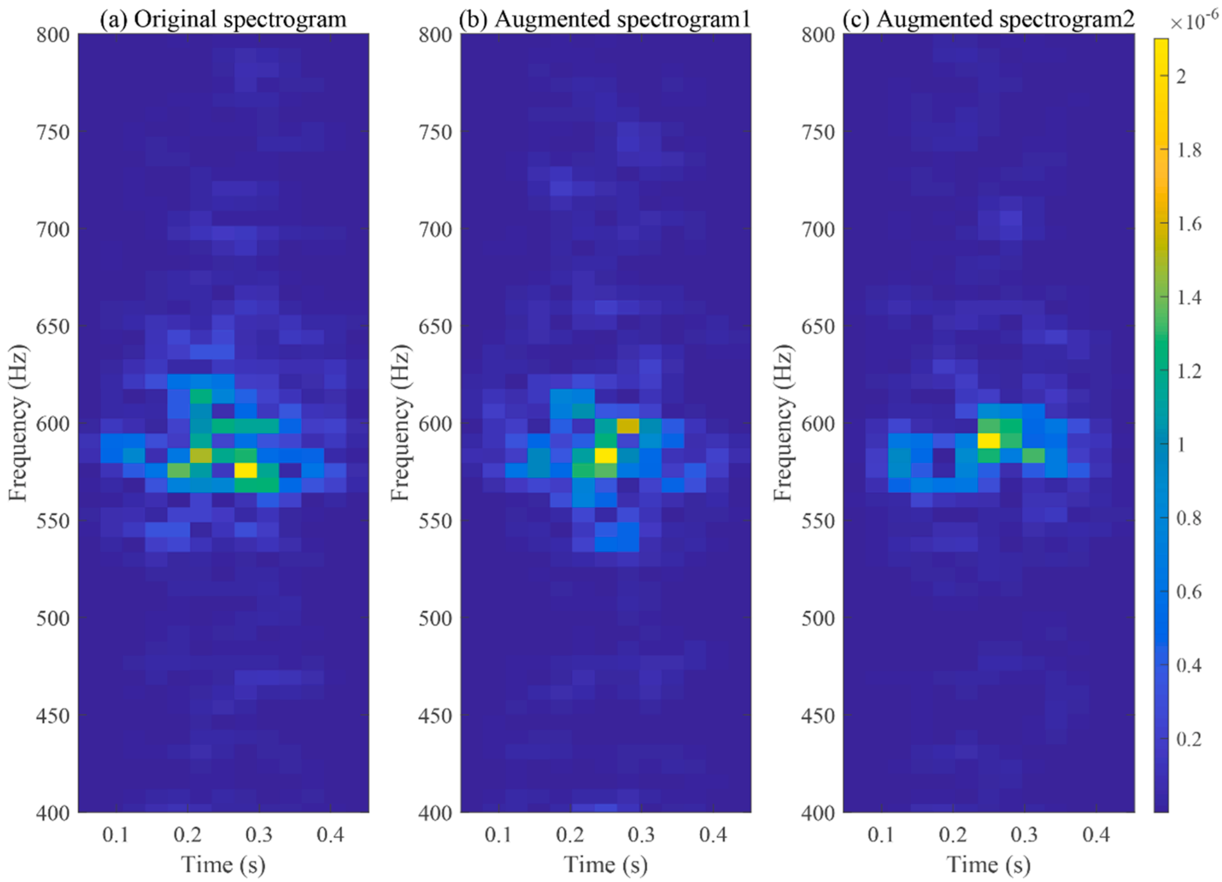

Fig. 7. Comparison of spectrograms between augmented data generated by IAAFT and the original signal.

compared to the baseline, showing a slight improvement over the gains achieved by IAAFT alone. In this context, IAAFT primarily contributes to enhancing data randomness, while masking compels the deep learning classifier to extract more global features from the already diverse dataset. The collaborative and mutually supportive role of IAAFT and masking is evident in the sequential manner.

Contrastingly, the combined data augmentation technique in the parallel manner surprisingly underperforms and is inferior to either IAAFT or masking alone. The rationale behind this lies in the fact that the two individual techniques enhance leak detection performance from distinct perspectives. Global features learned from partially masked spectrograms may not be applicable to IAAFT-generated data, creating a subtle antagonism in the parallel manner.

3.5. *Influence of data augmentation on transfer learning*

This section investigates whether data augmentation can further enhance the effectiveness of transfer learning. The study directly employs the combined data augmentation technique in the sequential manner that performed best on TFCNN-post-splitting. The model transferred in this study is PANN, with its hyper-parameters detailed in Table 1. Without data augmentation, PANN achieves an AUC, accuracy, sensitivity, and specificity of 0.810, 0.730, 0.747, and 0.713, respectively, serving as the baseline for the PANN model. The impact of data augmentation on PANN and the differences in data augmentation effects between PANN and TFCNN-post-splitting are illustrated in Fig. 9.

There is a notable gap between the two models' baseline, with PANN significantly outperforming TFCNN-post-splitting, highlighting the advantage of transfer learning. Furthermore, after applying data augmentation, PANN's performance improves further, albeit not as markedly as TFCNN-post-splitting. In addition, with an increase in augmented data volume, there is no significant improvement in the performance of PANN. This is primarily because the original data used for data augmentation in this study consisted of only 300 samples. The augmented data was generated based on features from the original data, which lacked diversity. Even with the optimal data augmentation techniques employed in this study, it is not possible to perfectly simulate acoustic signals from new scenarios (different pipe materials, diameters, and operational states). Therefore, there remains a significant disparity between the training data and the validation/testing data. In such cases, complex models like PANN are prone to overfitting on the training data, affecting their generalization ability (i.e., significantly reduced performance on the testing dataset). C. Zhang et al. (2022) also observed similar phenomena in their study of transfer learning-based acoustic leak detection. The results suggest that the effectiveness of data augmentation has its upper limit, making it necessary to collect more data.

After data augmentation, PANN's overall accuracy is slightly lower than TFCNN-post-splitting, likely due to the default classification threshold of 0.5 used in this study. Despite this, PANN's AUC is generally superior to TFCNN-post-splitting, indicating better overall performance across all threshold scenarios. Additionally, at a threshold of 0.5, PANN exhibits better sensitivity, i.e., it can identify more leak signals, whereas TFCNN-post-splitting excels in reducing false positives.

3.6. *Study limitations and future research*

The data augmentation techniques employed in this study contribute to improving acoustic leak detection performance, resulting in an accuracy increase of over 7%. However, considering the limited volume of original data used in this study, the efficacy of these techniques needs further validation on more extensive datasets in the future. It is also

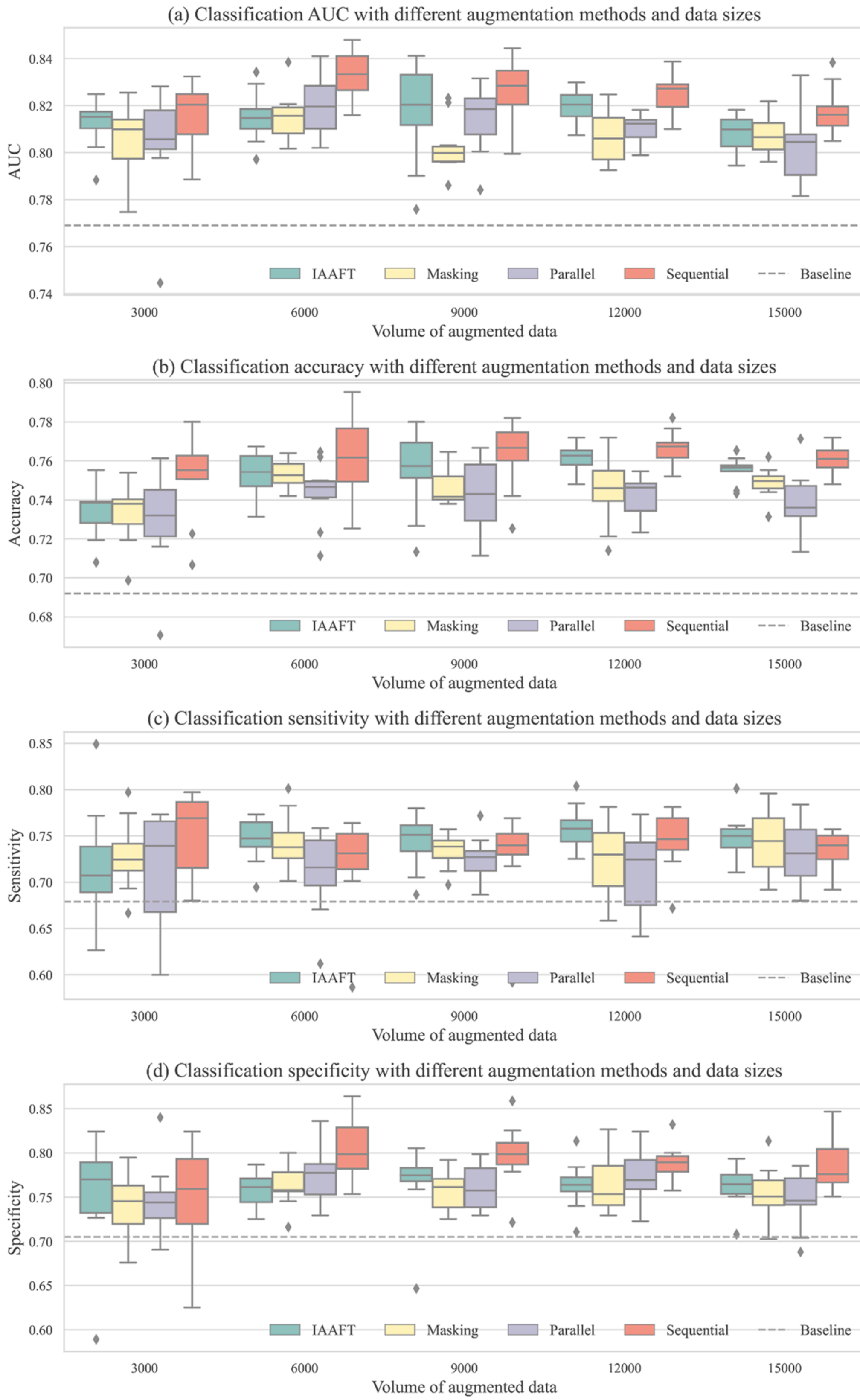

Fig. 8. Comparison of leak detection performance when using different combined data augmentation techniques.

noteworthy that despite the enhancement provided by data augmentation techniques in acoustic leak detection performance, the final performance metrics obtained still remain relatively low. The deep learning-based acoustic leak detection method has yet to reach the realm of high-performance practical applications. Mature application of AI technology requires multifaceted support. Future studies should prioritize the collection of more diverse data, encompassing various leak sizes, pipe materials, operational conditions, and more. After all, rich

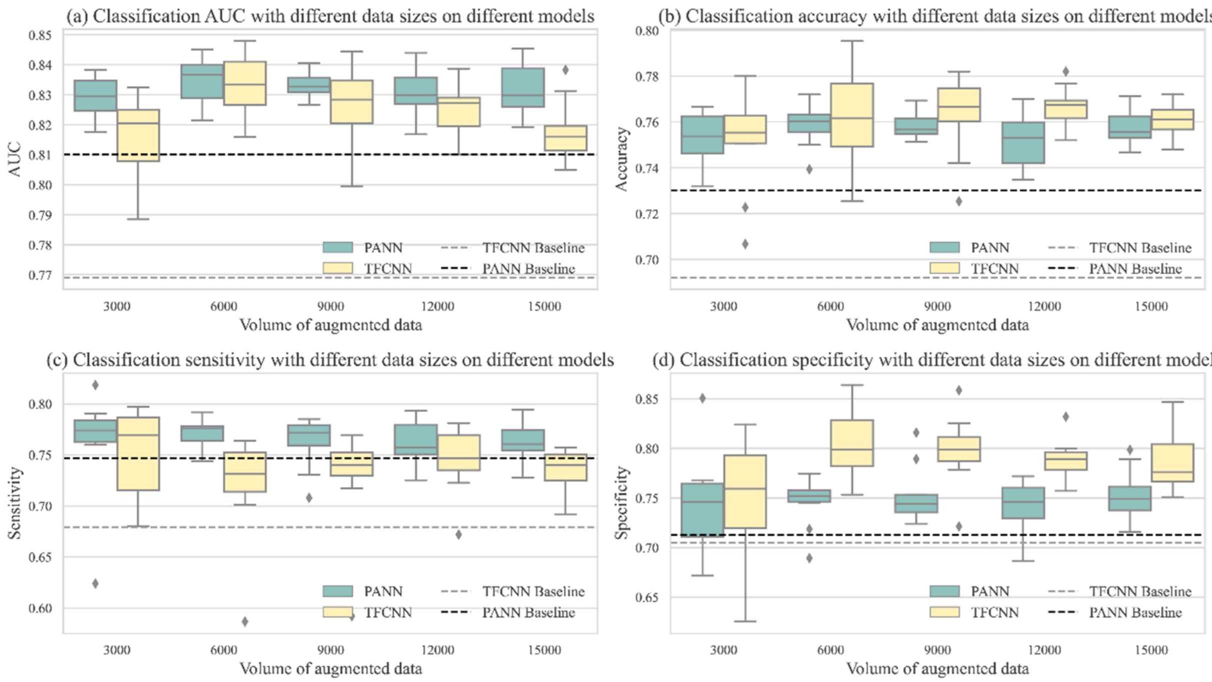

Fig. 9. Comparison of leak detection performance of PANN and TFCNN when using combined data augmentation technique.

raw data is the foundation for successful AI technology. In addition, the combined use of deep learning methods better suited for handling variable data and data augmentation techniques could be also crucial for further performance improvements.

Data-centric AI technologies, encompassing data augmentation techniques, are undergoing continuous evolution. Beyond the fundamental random transformation methods in time, frequency, and time-frequency domains, advanced time series data augmentation techniques incorporate sophisticated machine learning-based approaches, such as reinforcement learning and generative adversarial networks (Hataya et al., 2020; Wen et al., 2021). In this preliminary exploration of data augmentation for acoustic leak detection, the study focuses on basic random transformation approaches, namely jittering, scaling, warping, IAAFT, and masking. Future endeavors should delve into advanced techniques for a more comprehensive understanding.

4. Conclusions

This study explores the impact of data augmentation techniques on improving deep learning-based acoustic leak detection methods, aiming to advance AI-driven acoustic leak detection technology from a data-centric perspective towards more mature applications. After using CNN classifiers to identify augmented data generated by techniques such as jittering, scaling, warping, IAAFT, and masking, the following conclusions were drawn:

It is recommended to implement data augmentation before data splitting to avoid data leakage, which could lead to biased and overly optimistic results. Augmented data, being a random transformation of the original acoustic signals, retains certain similarities with the original data. If augmentation is performed before splitting, this similarity can lead the classifier to learn information during training that it should not have access to, introducing bias.

IAAFT and masking proved to be the most effective among the five proposed data augmentation techniques. IAAFT, significantly increasing both data volume and diversity while ensuring consistency with the labels of the original data (leak or non-leak), achieved the best results, improving accuracy by over 7%. Masking, by compelling the classifier to learn global features of the time-frequency spectrograms, also demonstrated notable performance enhancement. Data augmentation can also enhance the effectiveness of transfer learning, thereby creating better conditions for the application of complex deep learning models in acoustic leak detection.

The study recommends using IAAFT alone or sequentially applying IAAFT followed by masking. When IAAFT and masking are applied sequentially for data augmentation, they synergize, further enhancing leak detection performance. However, simultaneous use of both techniques results in some antagonistic effects, causing recognition performance to be inferior to either technique individually.

These conclusions provide practical insights for improving acoustic leak detection methods and also contribute to advancing the application of other AI-driven technologies in the area of WDSs.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT 3.5 in order to improve language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

CRediT authorship contribution statement

Yipeng Wu: Writing – review & editing, Writing – original draft, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Xingke Ma:** Validation, Investigation, Data curation, Software. **Guancheng Guo:** Writing – review & editing, Methodology, Investigation. **Tianlong Jia:** Software, Methodology. **Yujun Huang:** Visualization, Software. **Shuming Liu:** Writing – review & editing, Supervision, Resources. **Jingjing Fan:** Resources, Investigation. **Xue Wu:** Project administration, Investigation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Yipeng Wu reports financial support was provided by National

Natural Science Foundation of China. If there are other authors, they 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

The authors do not have permission to share data.

Acknowledgments

This work was financially supported by National Natural Science Foundation of China (Grant No. 52300119).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2024.121999.

References

Ahmad, S., Ahmad, Z., Kim, C., Kim, J., 2022. A method for pipeline leak detection based on acoustic imaging and deep learning. Sensors 22 (4), 1562.
Alqudah, R., Al-Mousa, A.A., Abu Hashyeh, Y., Alzaibaq, O.Z., 2023. A systemic comparison between using augmented data and synthetic data as means of enhancing wafermap defect classification. Comput. Ind. 145, 103809.
Bakhtawar, B., Zayed, T., 2021. Review of water leak detection and localization methods through hydrophone technology. J. Pipel. Syst. Eng. Pract. 12 (4), 3121002.
Bozkurt, C., Firat, M., Ates, A., 2022. Development of a new comprehensive framework for the evaluation of leak management components and practices. AQUA 71 (5), 642–663.
Chollet, F., 2023. Complete guide to transfer learning & fine-tuning in Keras.
Iwana, B.K., Uchida, S., 2021. An empirical survey of data augmentation for time series classification with neural networks. PLoS ONE 16 (7), e254841.
Cody, R.A., Dey, P., Narasimhan, S., 2020. Linear prediction for leak detection in water distribution networks. J. Pipel. Syst. Eng. Pract. 11 (1), 4019043.
El-Zahab, S., Zayed, T., 2019. Leak detection in water distribution networks: an introductory overview. Smart Water 4 (1), 5.
Fan, H., Tariq, S., Zayed, T., 2022. Acoustic leak detection approaches for water pipelines. Autom. Constr. 138, 104226.
Fares, A., Tijani, I.A., Rui, Z., Zayed, T., 2023. Leak detection in real water distribution networks based on acoustic emission and machine learning. Environ. Technol. 44 (25), 3850–3866.
Fox, S., Shepherd, W., Collins, R., Boxall, J., 2016. Experimental quantification of contaminant ingress into a buried leaking pipe during transient events. J Hydraul Eng 142 (1), 4015036.
Fu, G., Jin, Y., Sun, S., Yuan, Z., Butler, D., 2022. The role of deep learning in urban water management: a critical review. Water Res 223, 118973.
Guo, G., Yu, X., Liu, S., Ma, Z., Wu, Y., Xu, X., Wang, X., Smith, K., Wu, X., 2021. Leakage detection in water distribution systems based on time–frequency convolutional neural network. J. Water Resour. Plan. Manage.-ASCE 147 (2), 4020101.
Guo, G., Yu, X., Liu, S., Xu, X., Ma, Z., Wang, X., Huang, Y., Smith, K., 2020. Novel leakage detection and localization method based on line spectrum pair and cubic interpolation search. Water Resour. Manag. 34 (12), 3895–3911.
Hataya, R., Zdenek, J., Yoshizoe, K., Nakayama, H., 2020. Faster AutoAugment: Learning augmentation Strategies Using Backpropagation. Vedaldi, A., Bischof, H., Brox, T. and Frahm, J. (eds.), pp. 1–16, Springer International Publishing, Cham.
Heidenreich, N.B., Schindler, A., Sperlich, S., 2013. Bandwidth selection for kernel density estimation: a review of fully automatic selectors. Asta-Adv. Stat. Anal. 97 (4), 403–433.
Kang, J., Park, Y., Lee, J., Wang, S., Eom, D., 2018. Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems. IEEE Trans. Ind. Electron. 65 (5), 4279–4289.
Kong, Q., Cao, Y., Iqbal, T., Wang, Y., Wang, W., Plumbley, M.D., 2020. PANNs: large-scale pretrained audio neural networks for audio pattern recognition. IEEE/ACM Trans. Audio Speech Lang. Process. 28, 2880–2894.
Le Guennec, A., Malinowski, S., Tavenard, R., 2016. Data augmentation for time series classification using convolutional neural networks, Riva Del Garda, Italy.
Lee, T.E.K.M., Kuah, K., Leo, K.H., Sanei, S., Chew, E., Zhao, L., 2019. Surrogate rehabilitative time series data for image-based deep learning. In: 2019 27Th European Signal Processing Conference (Eusipco), pp. 1–5.
Li, S., Song, Y., Zhou, G., 2018. Leak detection of water distribution pipeline subject to failure of socket joint based on acoustic emission and pattern recognition. Measurement 115, 39–44.
Microsoft, 2021. Neural Network Intelligence (version v2.7).
Park, D.S., Chan, W., Zhang, Y., Chiu, C., Zoph, B., Cubuk, E.D., Le, Q.V., 2019. Specaugment: a simple data augmentation method for automatic speech recognition. Arxiv Preprint Arxiv:1904.08779.
Rashid, K.M., Louis, J., 2019. Times-series data augmentation and deep learning for construction equipment activity recognition. Adv. Eng. Inform. 42, 100944.
Ravichandran, T., Gavahi, K., Ponnambalam, K., Burtea, V., Mousavi, S.J., 2021. Ensemble-based machine learning approach for improved leak detection in water mains. J. Hydroinform. 23 (2), 307–323.
Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6 (1), 60.
Shukla, H., Piratla, K., 2020. Leakage detection in water pipelines using supervised classification of acceleration signals. Autom. Constr. 117, 103256.
Singh, P., 2023. Systematic review of data-centric approaches in artificial intelligence and machine learning. Data Sci. Manag. 6 (3), 144–157.
Sitaropoulos, K., Salamone, S., Sela, L., 2023. Frequency-based leak signature investigation using acoustic sensors in urban water distribution networks. Adv. Eng. Inform. 55, 101905.
Smith, K., Liu, S., Liu, Y., Guo, S., 2018. Can China reduce energy for water? A review of energy for urban water supply and wastewater treatment and suggestions for change. Renew. Sustain. Energy Rev. 91, 41–58.
Steven Eyobu, O., Han, D.S., 2018. Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network. Sensors 18 (9), 2892.
Tan, M., Le, Q., 2019. EfficientNet: rethinking model scaling for convolutional neural networks, pp. 6105–6114.
Ullah, N., Zahoor, A., Jong-Myon, K., 2023. Pipeline leakage detection using acoustic emission and machine learning algorithms. Sensors 23 (6), 3226.
Um, T.T., Pfister, F.M.J., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U., Kulić, D., 2017. Data Augmentation of Wearable Sensor Data for Parkinson's Disease Monitoring Using Convolutional Neural Networks. Association for Computing Machinery, Glasgow, UK, pp. 216–220.
Vanijjirattikhan, R., Khomsay, S., Kitbutrawat, N., Khomsay, K., Supakchukul, U., Udomsuk, S., Suwatthikul, J., Oumtrakul, N., Anusart, K., 2022. AI-based acoustic leak detection in water distribution systems. Results Eng 15, 100557.
Venema, V., Ament, F., Simmer, C., 2006. A Stochastic Iterative Amplitude Adjusted Fourier Transform algorithm with improved accuracy. Nonlin. Processes Geophys. 13 (3), 321–328.
Wen, Q., Sun, L., Yang, F., Song, X., Gao, J., Wang, X., Xu, H., 2021. Time series data augmentation for deep learning: a Survey. Int. Joint Conf. Artific. Intell. Org.
Wu, Y., Liu, S., Wu, X., Liu, Y., Guan, Y., 2016. Burst detection in district metering areas using a data driven clustering algorithm. Water Res 100, 28–37.
Wu, Y., Ma, X., Guo, G., Huang, Y., Liu, M., Liu, S., Zhang, J., Fan, J., 2023. Hybrid method for enhancing acoustic leak detection in water distribution systems: integration of handcrafted features and deep learning approaches. Process Saf. Environ. Protect. 177, 1366–1376.
Yu, T., Chen, X., Yan, W., Xu, Z., Ye, M., 2023. Leak detection in water distribution systems by classifying vibration signals. Mech. Syst. Signal Proc. 185, 109810.
Zha, D., Lai, K., Yang, F., Zou, N., Gao, H., Hu, X., 2023. Data-centric AI: Techniques and Future Perspectives. Association for Computing Machinery, Long Beach, CA, USA, pp. 5839–5840.
Zhang, C., Alexander, B.J., Stephens, M.L., Lambert, M.F., Gong, J., 2022a. A convolutional neural network for pipe crack and leak detection in smart water network. Struct. Health Monitor. 22 (1), 232–244.
Zhang, P., He, J., Huang, W., Yuan, Y., Wu, C., Zhang, J., Yuan, Y., Wang, P., Yang, B., Cheng, K., Cui, H., Zhang, L., 2022b. Ground vibration analysis of leak signals from buried liquid-filled pipes: an experimental investigation. Appl. Acoust. 200, 109054.
Zhang, P., He, J., Huang, W., Zhang, J., Yuan, Y., Chen, B., Yang, Z., Xiao, Y., Yuan, Y., Wu, C., Cui, H., Zhang, L., 2023. Water pipeline leak detection based on a pseudo-siamese convolutional neural network: integrating handcrafted features and deep representations. Water (Basel) 15 (6), 1088.
Zhang, X., Chen, N.C., Sheng, H., Ip, C., Yang, L., Chen, Y., Sang, Z., Tadesse, T., Lim, T., Rajabifard, A., Bueti, C., Zeng, L.L., Wardlow, B., Wang, S., Tang, S.Y., Xiong, Z., Li, D., Niyogi, D., 2019. Urban drought challenge to 2030 sustainable development goals. Sci. Total Environ. 693, 133536.
Zhu, J., Yang, M., Ren, Z.J., 2023. Machine learning in environmental research: common pitfalls and best practices. Environ. Sci. Technol. 57 (46), 17671–17689.