

Received 30 May 2024, accepted 14 June 2024, date of publication 17 June 2024, date of current version 24 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3416056



RESEARCH ARTICLE

Water Leakage Classification With Acceleration, Pressure, and Acoustic Data: Leveraging the Wavelet Scattering Transform, Unimodal Classifiers, and Late Fusion

ERICK AXEL MARTINEZ-RÍOS^{ID}, DAVID BARRIENTOS^{ID}, AND ROGELIO BUSTAMANTE^{ID}

School of Engineering and Sciences, Tecnológico de Monterrey, Mexico City 14380, Mexico

Corresponding author: Erick Axel Martinez-Ríos (a01331212@tec.mx)

The work of Erick Axel Martinez-Ríos was supported by Tecnológico de Monterrey and Consejo Nacional de Humanidades, Ciencias y Tecnologías under Grant CVU: 1010770.

ABSTRACT Early detection of water leakages is crucial due to their social, environmental, and economic impacts. In this regard, machine learning (ML) algorithms have been proposed in the literature to automatically detect leakages using vibration, pressure, and acoustic data. However, these data modalities are often used independently and under heterogeneous conditions. ML techniques typically involve feature extraction through the time and frequency domain, but the computation of these features varies among studies. On the other hand, deep learning (DL) techniques can automatically extract features from data but require large sample sizes and training times, and their complexity is higher than linear classifiers. This paper proposes the wavelet scattering transform (WST) as a feature extraction technique for hydrophone, vibration, and dynamic pressure data to compare their performance for classifying water leakages in looped and branched water networks. Late Fusion (LF) was also used to assess the effectiveness of simultaneously employing vibration, hydrophone, and dynamic pressure data to classify water leakages. The results indicate that the WST of accelerometer data and a support vector machine (SVM) performed best in classifying water leakages compared to using the WST of dynamic pressure and hydrophone data, with an accuracy of 89.5833% for the looped water network and 96.6667% for the branched water network. The LF model generated by combining the predictions of the SVMs of each data modality achieved a 93.33% accuracy for the looped water network. In contrast, the LF model of the branched water network attained an accuracy of 95.833%.

INDEX TERMS Water leak classification, wavelet scattering transform, feature extraction, accelerometer, pressure, hydrophone, late fusion.

I. INTRODUCTION

Water consumption is an essential human right and crucial for the growth of human societies. Nevertheless, water pipeline leakages are a critical challenge affecting water consumption and distribution. Water pipeline leakages can be caused by pipeline deterioration, temperature, pressure, human damage, and geological changes [1], [2]. Most of this water loss goes unnoticed since water distribution networks are underground,

or the leakages are relatively small [3], [4]. The above could have several social effects, such as water contamination, traffic delay, and water scarcity [5]. According to the World Bank, water leakages produce about 48.6 billion cubic meters of water losses annually. The above also produces economic losses of about 14.6 billion dollars [6]. Hence, it is crucial to promptly and precisely detect and attend to water leakages in water distribution networks.

Nevertheless, water leakage detection is a challenging task influenced by consumer demand's dynamic and uncertain behavior. The above implies that not all pressure drops or

The associate editor coordinating the review of this manuscript and approving it for publication was Andrea De Marcellis^{ID}.

flow increases can be related to water leaks [7]. Additionally, the sensors that monitor water distribution systems require frequent calibration and are susceptible to noise, exacerbating the leakage detection task [8]. In this regard, efficient detection and classification of water leakages across water distribution networks are of utmost importance and have led to the development of automated water leakage detection systems [9].

Water leakage detection methods can be categorized into two main groups: hardware-based and software-based methods [10]. Hardware-based leak detection utilizes real-time monitoring devices to monitor the pipeline infrastructure. These devices include hydrophone, fiber optic, thermographic, ultrasonic, and electromagnetic devices. Nonetheless, hardware-based methods are expensive, time-consuming, and inappropriate for long-term monitoring [11].

On the other hand, two main approaches exist within software-based methods: model-based and data-driven. Model-based techniques utilize the estimation of the pipeline parameters and numerical modeling to produce models that can detect water leakages. Nevertheless, estimating the pipeline parameters requires the assumption of an intact or original condition in the pipe, may require lengthy processing times, and their calibration could be challenging due to measurement and modeling uncertainties [12], [13].

The availability of collecting a considerable amount of historical data related to water leakages has facilitated the utilization of data-driven techniques for detecting and locating leaks [5]. Furthermore, contrary to model-based methods, data-driven techniques do not require a deep understanding of the water network characteristics [14]. Machine learning techniques are one of the data-driven techniques employed to fit classification models that consider different data modalities as input, including flow, pressure, acoustic, and vibration data to detect water leakages [15], [16], [17].

Despite the widespread use of machine learning algorithms for generating models that allow water leakage detection, the trained models have used only one type of data modality for training the machine learning algorithms [18], [19], [20], [21], [22], [23], [24], [25], [26]. In certain studies, flow and pressure data have been used simultaneously as input for water leakage detection [27], [28], [29], [30], [31]; however, comparing different data modalities to perform water leakage detection based on machine learning still needs to be explored. The above is crucial to determine which sensor types are the most adequate to detect water leakages. Moreover, the studies that have used different data modalities are difficult to compare due to the heterogeneity between the scenarios and datasets the authors have employed. Thus, it is critical to compare the effectiveness of different data modalities for water leakage detection based on machine learning under equal conditions.

Another challenge that involves using vibration, acoustic, pressure, or flow sensor data is to compute adequate features that allow the discrimination between the presence of

leakages and their absence. While deep learning techniques such as convolutional neural networks (CNNs) can automatically extract features from the input data, they often require large sample sizes to avoid overfitting [32], [33]. At the same time, gathering quality data could be unaffordable and challenging [34]. Furthermore, training deep learning techniques requires significant computational resources and large training times [15]. In addition, the architecture of deep neural networks is often complex compared to that of linear classifiers, which hinders their interpretability [35].

On the other hand, the manual extraction of features from signals or time series data requires domain knowledge to compute appropriate features that allow the discrimination between water leakages and no leakages [36], [37]. Several features could be computed based on the representation in which the data is analyzed, such as the time domain (e.g., root mean square, zero crossing rate, entropy, and standard deviation), frequency domain (e.g., power spectral density, and peak frequency), and time-frequency domain (e.g., short-time Fourier transform and wavelet transform) [38], [39], [40], [41], [42]. This diverse set of signal representations brings out the challenge of computing features in an adequate representation that allows the discrimination between signals associated with the presence or absence of water leakages in water distribution networks regardless of the type of sensor used to perform the leak detection.

Moreover, the literature has mainly addressed the issue of water leak detection in pipelines as a binary classification problem [18], [23], [37], [43], [44]. The above means that the proposed methods can only identify whether there is a water leak or not. Thus, there needs to be more exploration into how to classify the different types of water leaks that can occur in water pipelines.

This paper compares different data modalities for water leakage classification by employing machine learning techniques under equal conditions. The compared data modalities include acceleration, acoustic, and dynamic pressure data. Furthermore, it is proposed to use the Wavelet Scattering Transform (WST) to extract features from the time series data for each of the considered modalities. The WST is a signal representation technique that creates a signal representation invariant to translations and stable to small-time warping deformations [45]. The above is done to minimize the within-class variance and, in this way, produce discriminative features that allow water leakage classification from acceleration, acoustic, and dynamic pressure data without requiring manual feature extraction and deep learning. Moreover, Late Fusion (LF) was used to consider the effect of each data modality for water leak classification and compare its performance with the models trained with only one type of data modality. Consequently, this study is concerned with comparing the use of different data modalities to perform water leakage classification based on machine learning and using the scattering transform as a feature extraction technique to avoid deep learning techniques and manual feature extraction. Fig. 1 shows an overview of the

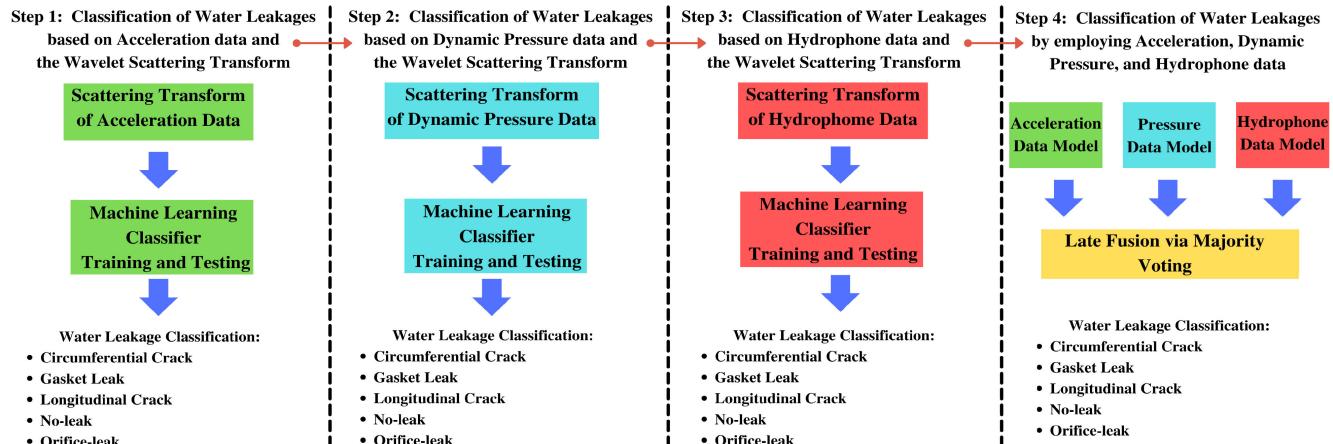


FIGURE 1. Overview of the methodology followed in this study for water leakage classification. The process involves four steps to identify water leakages using machine learning. The first step is to use the WST on accelerometer data to extract features to train a machine learning classifier. The exact process is repeated for dynamic pressure data in the second step. The third step involves extracting features from hydrophone data using the WST to train a machine-learning classifier. Finally, in the fourth step, the previously trained models for each data modality are utilized, and their predictions are combined via LF using a majority voting mechanism.

methodology used in this study. The main contributions of this study are listed as follows:

- The WST is used as a feature extraction technique for accelerometer/vibration, hydrophone, and dynamic pressure data to classify water leakages.
- Classification of four water leakage types via machine learning and the WST that include orifice leak, longitudinal and circumferential cracks leaks, gasket leak, as well as no-leaks in looped and branched networks.
- A comparison of using accelerometer, hydrophone, and dynamic pressure data for water leakage classification based on machine learning in the looped and branched water networks.
- An evaluation of LF to perform multimodal water leakage classification by employing a majority voting mechanism and the predictions of the unimodal models trained by using the WST of accelerometer/vibration, dynamic pressure, and hydrophone data.

The rest of this study is organized as follows. Section II presents a brief literature review of studies that have utilized machine learning techniques for water leakage detection, emphasizing the data modality, feature extraction technique, and machine learning techniques employed in each study. Section III shows the dataset used to develop the present study and an overview of the theoretical background of the methods and algorithms used in this work. Section IV presents the results of the proposed methodology, while Section V shows the discussions. The study's limitations are presented in Section VI. Finally, Section VII presents the conclusions and future research directions.

II. LITERATURE REVIEW

The problem of water leakage detection has been analyzed and tackled by employing different types of data modalities and data-driven techniques, where machine learning techniques have picked up a certain degree of attention. The available studies can be classified based on the sensor or

input data the authors used to train the machine-learning technique. Generally, the input data includes acoustic emission, vibration, flow, and pressure data. According to Islam et al. [15], these sensors are frequently used since they are low-cost and easy to install. This section provides an overview of the proposed water leakage detection methods based on machine learning. The reviewed articles have been categorized according to the type of input data that each study analyzed. If more than one data modality was used in the study, it was categorized in a section dedicated to multimodal data. Table 1 summarizes the related research presented in this section.

A. FLOW SENSOR DATA FOR WATER LEAKAGE DETECTION BASED ON MACHINE LEARNING

Flow sensors have been employed to monitor and detect water leakages in distribution networks. For example, Fereidooni et al. [19] utilized flow sensor data to develop a hybrid model for water leakage detection and localization by employing hydraulic equations as feature extractors. The generated features were used to train a support vector machine (SVM), a k-nearest neighbor (KNN), a decision tree (DT), a random forest, and a Bayesian network. From these classifiers, the Bayesian network outperforms the other methods regarding the F1-score with a value of 95.775%. The water network scenarios were simulated using the Water GEMS simulator. One disadvantage of this approach is that calculating hydraulic features requires knowing the pipeline parameters, such as the pipe's length, diameter, and roughness, which may not be accessible or available. Moreover, flow sensor data collection has disadvantages, such as an expensive and in-pipe installation [15].

B. VIBRATION DATA FOR WATER LEAKAGE DETECTION BASED ON MACHINE LEARNING

Vibration data has become a widely applied sensor for detecting water leakages in water pipelines. Recently, a study

TABLE 1. Summary of the related research focused on water leakage detection based on machine learning.

Author	Year	Signals and Sensors	System	Features	Machine Learning Technique	Performance%
Liu et al. [1]	2019	Acoustic data	Aluminum-plastic composite pipe	Intrinsic mode function, and approximate entropy	SVM	Accuracy: 98
Guo et al. [18]	2020	Piezoelectric accelerometer	Cheng Du city water pipeline network	Short-Time Fourier Transform Spectrogram	CNN	Accuracy: 99
Fereidooni et al. [19]	2020	Flow sensors	Simulation of Water Network	Hydraulic Equations	Bayesian Network	Accuracy: 93.80375 Precision: 96.27 Recall: 95.9375 F1-score: 95.775
Ravichandran et al. [46]	2021	Acoustic data	Urban networks of the United States and Canada	Time domain Features and power spectral density	Gradient Boosting Tree	Accuracy: 93.5
Hu et al. [31]	2021	Pressure and Flow Data	Hydraulic simulation performed on EPANET	Residuals of pressure and flow data	MFCN	Mean Per-Class Error: 0
Mashhadi et al. [29]	2021	Pressure and Flow Data	Water distribution system of Lille University	Water flow and pressure measurements	DT	Accuracy: 98
Bohorquez et al. [12]	2022	Transient head pressure traces	Single laboratory water pipeline	Raw transient pressure head traces	1D CNN	Error: 0.59
Vanijjirattikhan et al. [20]	2022	Acoustic rod/Microphone	Water supply network of Bangkok	Frequency domain Features	Deep Neural Network	Accuracy: 94.89
Kammoun et al. [27]	2022	Pressure and Flow Data	LeakDB dataset	Raw pressure and flow data	SVM	Accuracy: range from 30 to 100
Tariq et al. [47]	2022	Accelerometers	Water network composed of metallic and non-metallic pipes	Standard deviation and monitoring index	Random Forest	Accuracy for metal pipes: 100 Accuracy for non-metal pipes: 94.93
Shen et al. [21]	2022	Acoustic Emission	Water distribution systems of Hangzhou, Shaoxing, and Shanghai.	Frequency domain Features	Random Forest	Recall: 100
Yu et al. [22]	2023	Accelerometers	Water distribution networks of cities of China	Spectrogram	SqueezeNet	Accuracy: 95.15
Sousa et al. [23]	2023	Pressure	District Meter Areas of Stockholm, Sweden	Features obtained via the canonical discriminant function	Linear Vector Quantization	F1-score: 91.73
Fares et al. [24]	2023	Acoustic Emission	Water Distribution Network of Hong Kong	Time and Frequency domain Features	Deep Neural Network	Accuracy: 87.2 F1-score: 79.5
Saravanabalaji et al. [25]	2023	Acoustic Emission	Laboratory water pipeline	Frequency domain Features	Hybrid Machine Learning Model	Accuracy: 99.3
Ullah et al. [26]	2023	Acoustic Emission	Laboratory pipeline	Frequency domain and time domain Features	Neural Network	Accuracy: 100
Sahin et al. [28]	2023	Pressure and Flow Data	Laboratory water pipeline	Flow and Pressure data measurements	Graph Convolutional Network	Accuracy: 94
Tyagi et al. [30]	2023	Pressure and Flow Data	Three water distribution networks: Hanoi, Net3, and C-town	Flow and Pressure data measurements at network nodes	Multinomial Logistic Regression	Accuracy: range from 34 to 92
Lee et al. [44]	2023	Vibration Data	Water network composed of metallic and non-metallic pipes	Time-Frequency domain Features	XGBoost	Accuracy: 99.79

by Yu et al. [22] utilized vibration data to detect water leakages in actual water distribution pipelines. The authors collected vibration data through piezoelectric accelerometers from water distribution networks in various cities of China. They compared two approaches to detect water leakages. The first approach involved extracting time and frequency domain features from the vibration data to train an SVM, a DT, and a KNN. The second approach involved extracting time-frequency images of the vibration data using the Short-Time Fourier transform, which was then used to fine-tune the SqueezeNet. The results indicated that SqueezeNet performed better than other tested machine learning methods. Similarly, Guo et al. [18] analyzed the performance of employing time-frequency representations and CNNs by applying the Short-Time Fourier transform to piezoelectric accelerometer signals associated with water leakages underneath distinct signal-to-noise ratio conditions. The authors reported a mean accuracy of 99% for different time-frequency resolutions.

In addition, Lee and Kim [44] utilized time and frequency domain features from vibration sensors to perform water leakage detection. Lee's study compared different machine learning classifiers such as KNN, DT, random forest, gradient boosting, and XGboost; nonetheless, XGboost achieved the best performance with an accuracy of 99.79%. Besides,

Tariq et al. [47] explored using cost-effective microelectromechanical systems-based accelerometers to detect water leakages in metal and non-metal pipes. The authors suggested a method that uses the standard deviation and the monitoring index for feature extraction of accelerometer data. The features were then used to train machine learning algorithms like the KNN, DT, random forest, and Adaboost. The random forest algorithm achieved an accuracy of 100% for metal pipes and 94.93% for non-metal pipes.

As can be appreciated, vibration data is frequently employed to detect water leakages. However, this method has limitations, such as a short range and contamination by environmental vibrations [15]. It can be appreciated that time-frequency representations, along with CNNs, are often used to process the vibration data to detect water leakages. Nonetheless, the computation of time-frequency representations is time-consuming, as in the case of training CNNs [48], [49].

C. PRESSURE DATA FOR WATER LEAKAGE DETECTION BASED ON MACHINE LEARNING

Pressure data is another data modality employed to develop water leakage detection systems. For example, Sousa et al. [23] analyzed pressure data from pumps of

district meter areas of Stockholm, Sweden. To extract features from the time series pressure data, the authors employed a canonical discriminant function; these features were used to train a linear vector quantization algorithm that achieved a classification rate of 93.98%. On the other hand, Bohorquez et al. [12] proposed using a 1D CNN to detect water leaks by using pressure transient heads and enhancing their performance with different noise intensities. Despite the above, pressure data have weaknesses, such as the lack of capacity to detect small leaks and expensive installation [15].

D. ACOUSTIC DATA FOR WATER LEAKAGE DETECTION BASED ON MACHINE LEARNING

Acoustic data has been extensively used to detect water leakages. For example, the work of Fares et al. [24] proposed using acoustic emission data obtained from wireless noise loggers to detect water leakages. Fares employed time and frequency domain features computed from the acoustic emission signals. The computed features included skewness, kurtosis, peak amplitude, crest factor, peak frequency, frequency spread, and frequency centroid. The features were used to train tree-based methods, artificial neural networks, SVMs, logistic regression, Naïve Bayes, and rule induction. From these methods, the SVM and the artificial neural network achieved the best performance in terms of accuracy. Similarly, Shen and Cheng [21] suggested using acoustic emission signals to detect water leaks in water distribution systems by analyzing on-site data to account for the more significant interference and randomness to which signals are susceptible compared to laboratory test conditions. The features calculated from the acoustic emission signal were the main frequency, spectral flatness, roll-off rate, and mel-frequency cepstral coefficients. The authors employed tree-based methods to perform the detection, such as Adaboost, random forest, and DTs, with the Adaboost achieving the best performance in terms of recall.

Saravanabalaji et al. [25] collected acoustic emission signals and processed them using the Fourier transform to extract features for water leak detection. The Fourier representation of the acoustic signals was processed via principal component analysis to reduce the dimensionality of the feature vector, which was used to train a hybrid model composed of logistic regression, random forest, Adaboost, XGboost, and a linear and radial basis function (RBF) SVMs. Likewise, Ullah et al. [26] employed machine learning algorithms to detect water leaks from different pinhole-sized leaks via acoustic emission data. Time and frequency domain features were computed from the acoustic emission data to train a neural network, DT, random forest, and KNN. As in previous studies, the acoustic emission data was used to extract features such as kurtosis, skewness, mean value, mean square, root mean square, peak value, standard deviation, entropy, and frequency spectrum. According to the author's results, the four machine learning techniques achieved an accuracy of 99%.

Acoustic rod/microphone data have also been utilized for water leak detection. Vanijjirattikhan et al. [20] present a system for pinpointing leaking pipes based on microphone data. The acoustic data was used as input to train and compare an SVM, a deep neural network, and a CNN, with the deep neural network, achieves the best performance. According to the authors, the deep neural network achieved the best performance in terms of accuracy with a value above 90%. Liu et al. [1] employed acoustic sensors to acquire signals related to water pipeline leakages. The acoustic data generated a feature set through intrinsic mode function, approximate entropy, and principal component analysis. The generated features were used to train an SVM. The proposed method achieved an accuracy of 98%, according to Liu's study. Lastly, Ravichandran et al. [46] computed time domain features and the power spectral density of acoustic data to generate features to train a multi-strategy ensemble learning using a gradient boosting tree. Despite the extensive use of acoustic data in water leak detection presented above, background noise may degrade its performance [18].

E. MULTIMODAL DATA FOR WATER LEAKAGE DETECTION BASED ON MACHINE LEARNING

Particular works have also explored different data modalities to detect and locate water leakages. For instance, Mashhadi et al. [29] utilized flow and pressure data to compare the performance of six machine learning methods to detect leakages in water distribution networks. The compared methods were a logistic regression, DT, random forest, principal component analysis, k-means, and artificial neural networks, where the logistic regression and random forest achieved the highest performance by employing flow and pressure data as input separately. Nonetheless, only the DT was tested when simultaneously inputting the flow and pressure data. Similarly, Kammoun et al. [27] proposed using flow and pressure data for leakage detection and localization in water distribution networks. Kammoun's study compared an artificial neural network, a KNN, an SVM, and a logistic regression. However, Kammoun's study did not consider the flow and pressure data simultaneously as input for training and testing the machine learning techniques. Şahin and Yüce [28] analyzed flow and pressure data to train a Graph CNN and an SVM for water leakage prediction. According to Şahin's study, the accuracy obtained by the Graph CNN was 95%.

Likewise, Tyagi et al. [30] employed pressure and flow sensor data to train a two-stage model that fits a linear regression model to calculate the residual error of pressure heads in the water network. Subsequently, the predicted error distributions were used as input to train a multinomial logistic regression to detect the presence of a water leak. Hu et al. [31] proposed a density-based spatial clustering of applications with noise (DBSCAN) and a multiscale fully convolutional network (MFCN) for water leak detection. The DBSCAN was used to reduce the number of learning labels by dividing the water network into zones, which reduced the number

of output values of the MFCN. The pressure and flow data residuals were used as input into the MFCN. Hu's study used the Mean Per-Class Error as an evaluation metric where a zero value was reported for the MFCN.

F. LITERATURE REVIEW ANALYSIS

This literature review shows that different methodologies based on machine learning techniques have been used for water leakage detection. From the revised studies, acoustic data is the most frequently utilized data modality. Nevertheless, the authors have often considered only one type of data modality, such as acoustic data, vibration data, pressure or flow sensor data, to train the machine learning algorithms. Thus, there is an opportunity to compare and test the performance of different data modalities for water leakage detection in homogenous conditions since the data modalities have been tested on different water distribution systems or scenarios, as presented in Table 1. Hence, comparing different data modalities under equal conditions is crucial to determine the most appropriate sensor to train the classifier. Furthermore, fusion techniques to consider the effect of the different data modalities in water leakage detection performance based on machine learning remain to be explored extensively. Few works, such as those presented in [27], [28], [29], [30], and [31] have considered more than one data modality to detect water leakages based on machine learning. Nevertheless, the combined data modalities typically include flow and pressure data rather than acoustic or vibration data. Some studies have used multiple data types independently (i.e., training a model for each data or sensor type) without considering using them as input into the models simultaneously [27], [29].

Additionally, regardless of the data modality used as input, the research has focused on using deep learning algorithms, such as CNNs, to perform water leakage detection [50]. Nevertheless, this approach requires high computational power and a larger sample size, and their structure tends to be more significant than linear classifiers [51]. Still, deep learning algorithms did not require a manual feature extraction process. On the other hand, the computation of features commonly includes using time, frequency, and time-frequency domain representations. Nonetheless, a standardized set of features is not available, and a comparison of the performance of the different signal representation techniques in which the features can be computed is not available, which also leads to guessing the features that are going to be extracted from the sensor data to train the machine learning technique [52]. Hence, there is an opportunity to explore using signal representations that are both translation invariant and stable to deformations as feature extraction methods to avoid deep learning and manual feature extraction via the time domain and frequency domain.

Most of the literature has been concerned with detecting the presence of water leakages; this problem is tackled as a binary classification problem [18], [19], [23], [25], [43], [44],

[46]. Nevertheless, the classification of the water leak types, aside from only detecting the presence of the leakage, is an aspect that needs to be explored in the existing literature. The above implies determining which type of leakage is present in the water pipeline, such as circumferential, longitudinal, orifice, and gasket leaks.

III. MATERIALS AND METHODS

This section provides an overview of the dataset used in the present study and the theoretical background of the techniques utilized in this work. The preprocessing steps and procedures undertaken to conduct the study are outlined below.

A. DATASET

The dataset used in this study was published by Aghashahi et al. [53] and can be downloaded from the Mendeley open repository of Aghashahi et al. available in [54]. The dataset comprises 280 sensor measurements for water leakage classification and localization in water distribution networks. Three sensor measurements were considered: accelerometer, dynamic pressure, and hydrophone sensors. Four types of leaks were considered: orifice leaks, longitudinal and circumferential crack leaks, and gasket leaks. Additionally, no-leak signal measurements were recorded. The response of the three sensor types was recorded while one of the leakages was induced in the water network. Two network topologies were considered while generating the dataset looped and branched. The signal measurements lasted 30 seconds, while the sampling rate was 51.2 kS/s/ch for the dynamic pressure and accelerometer sensors and 8 kHz for the hydrophone data. Figs. 2(a) and 2(b) illustrate the water networks used to collect the leakage data. The supply line for the water network consists of a storage tank, a 25.4 mm Neptune MACH 10 ultrasonic meter with a 0.0038 L resolution (labeled as M1), a gate valve, a fixed-speed centrifugal pump model Goulds 1MC1G1A0, and a check valve. Moreover, PVC pipes were used to construct the networks.

As shown in Fig. 2, two hydrophones, two accelerometers, and two dynamic pressure sensors were placed along the water pipeline testbed. The accelerometer sensors measured the vibration of the water pipes, while the hydrophone and dynamics pressure sensors measured the sound and pressure variations inside them. The accelerometer sensors were located on the branches of two tee connections. Hydrophones require direct contact with water, typically mounted at the top or bottom of fire hydrants. Two fire hydrants were simulated and placed as far as possible from each other and symmetrical to the leak position. The dynamic pressure sensors were placed at the end of the supply line and the farthest point corner from the entry point. The different leak types were generated in the center of the water network testbed, as shown in Fig. 2 [53].

As previously mentioned, 280 samples were collected that were obtained by measuring the system's response

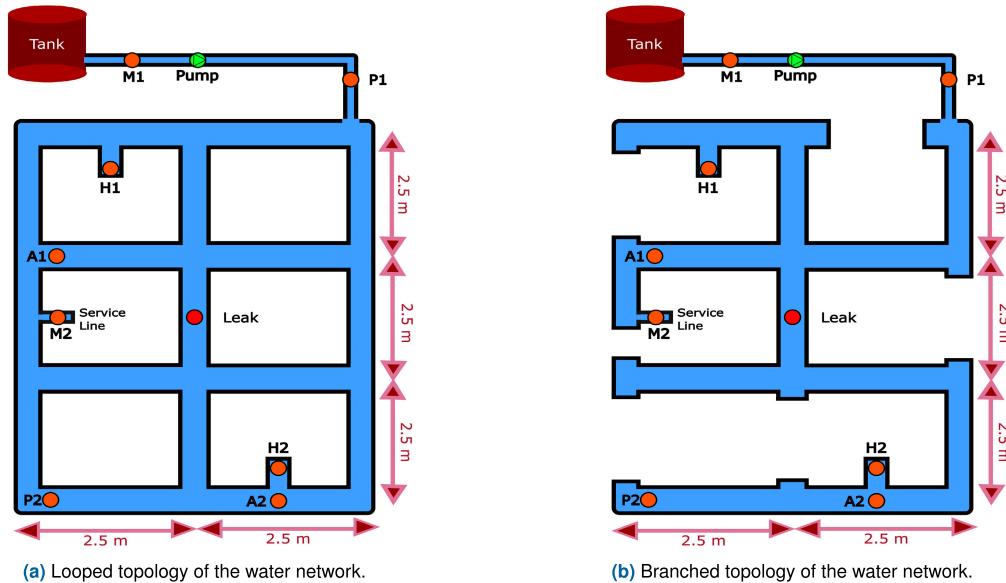


FIGURE 2. Illustration of the water network topologies of the dataset. a) Shows a schematic representation of the looped network where the letter A denotes the position of the accelerometer sensors, the letter H denotes the position of the hydrophone sensors, the letter P denotes the position of the pressure sensors, and the letter M denotes the position of flow meters. b) Shows a schematic representation of the branched network where the letter A denotes the position of the accelerometer sensors, the letter H denotes the position of the hydrophone sensors, the letter P denotes the position of the pressure sensors, and the letter M denotes the position of flow meters.

to variations in the topology of the network (i.e., looped and branched), water leak type, background condition, and sensors (i.e., hydrophone, accelerometer, and dynamic pressure). A total of 140 signals were recorded for each water network topology. On the other hand, background conditions refer to a combination of water demand and background noise. The service line of the water network (see Fig. 2) served to simulate water demand conditions. These water demands were 0, 0.18, and 0.47 L/s. A transient water demand condition was also generated by abruptly changing the demand from 0.47 L/s to 0 L/s. The background noise was only considered for the hydrophone data and was generated by simultaneously playing traffic and chainsaw sounds while recording the hydrophone data.

Considering the above, the response of each sensor was recorded four times for each water leak type by considering the 0, 0.18, 0.47 L/s, and transient water demands. Since there are two sensors for the hydrophone, accelerometer, and dynamic pressure data, the number of measurements associated with a single sensor type is eight. Furthermore, five classes are associated with each leak type, leading to 40 measurements associated with one sensor and network type. That is, eight measurements were associated with a gasket leak, eight were for an orifice leak, eight were for a longitudinal crack, eight were for a circumferential crack, and eight were related to no leaks. As a result, the number of signals linked to each water leak is balanced for each sensor type.

The only exception to this rule is the hydrophone data, which includes more than 40 measurements for a single

sensor type. The above is due to the inclusion of four additional measurements for each water leak type, where the hydrophone data was collected without background noise. These measurements were obtained by assessing the response of the two hydrophone sensors under background conditions of no demand and transient demand. The hydrophone measurements without background noise were excluded from this study to maintain balance in the number of measurements across each data modality and leak type.

Table 2 summarizes the dataset characteristics such as the sample size, sensors model and type, data acquisition devices, software, leak types, and background demand conditions. It is important to note that the number of signals associated with each water leak type was the same for both water network topologies. The above implies that the dataset is balanced for each class or water leak type considered during the data collection. Please refer to the work of Aghashahi et al. [53] for further details related to the data collection.

It is important to note that the response of the vibration, hydrophone, and dynamic pressure data may differ depending on the dynamics of the water network topology. To account for this, a separate model was fitted for each water network topology, one for the looped water network and another for the branched water network. As a result, each of the selected machine-learning techniques was used twice to fit a model for each water network topology.

B. WAVELET SCATTERING TRANSFORM

The WST was first presented and discussed in the seminal work of Mallat presented in [55]. The main goal of the WST is

TABLE 2. Summary of the dataset's main characteristics [53]. A1 and A2 refer to the accelerometer sensors 1 and 2, respectively. P1 and P2 refer to the dynamic pressure sensors 1 and 2, respectively. H1 and H2 refer to the hydrophone sensors 1 and 2, respectively.

	Dataset characteristics
Sample Size	280
Accelerometer model and quantity	Model: PCB 333B50 Quantity: 2
Hydrophone model and quantity	Model: Aquarian H2c Quantity: 2
Dynamic pressure sensor model and quantity	Model: PCB 102B16 Quantity: 2
Accelerometer units and Sampling Frequency	Units: $\frac{m}{s^2}$ Sampling Frequency: 51.2 kS/s/ch
Hydrophone units and Sampling Frequency	Units: Pascal (PA) Sampling Frequency: 8kHz
Dynamics pressure sensor units and Sampling Frequency	Units: Volts (V) and decibels (dB) via 20 log(10 V). Sampling Frequency: 51.2 kS/s/ch
Data acquisition devices	NI-9234 for accelerometer and dynamic pressure sensors. ZOOM UAC-2 audio converter for hydrophone data.
Software	LabVIEW NXG 5.1 for accelerometer and dynamic pressure data. Audacity 3.0.5 for hydrophone data.
Leak Types and Labels	Orifice Leak Longitudinal Leak Circumferential Leak Gasket Leak No Leak
Number of signals per network topology	140 signals for Looped network 140 signals for Branched network
Background demand conditions	No demand (0 L/s) 0.18 L/s 0.47 L/s Transient - demand abruptly changed from 0.47 L/s to 0 L/s
Background demand conditions considered to record the response of the system for each sensor type, and leak type	1 measurement of the response of the sensor to 0 L/s demand 1 measurement of the response of the sensor to 0.18 L/s demand 1 measurement of the response of the sensor to 0.47 L/s demand 1 measurement of the response of the sensor to transient demand
Number of signals for each leak type related to dynamic pressure data by considering the two available sensors	8 (4 related to sensor P1 and 4 to sensor P2)
Number of signals for each leak type related to hydrophone data by considering the two available sensors	12 (6 related to sensor H1 and 6 to sensor H2) Note: There are two additional sensor measurements for each hydrophone sensor where the data was recorded without background noise by considering as background condition the no demand and transient demand.
Number of signals for each leak type related to accelerometer data by considering the two available sensors	8 (4 related to sensor A1 and 4 to sensor A2)
Number of signal measurements associated with accelerometer data for each water network	40
Number of signal measurements associated with dynamic pressure data for each water network	40
Number of signal measurements associated with hydrophone data for each water network	60

to create a translation-invariant representation of a 1D signal or image that is stable to small deformations. The primary purpose of creating invariance to translations and stability to small deformations is to reduce the within-class variance of the set of signals to be classified. For the case of 1D signals, Andén and Mallat [45] define stability to deformations as a Lipschitz continuity condition as expressed in (1):

$$\|\Phi(x) - \Phi(x_\tau)\| \leq C \sup_t |\tau'(t)| \|x\| \quad (1)$$

where $C > 0$, x represents the original signal $x(t)$, x_τ represents the deformed version of the signal $x(t)$, $\tau(t)$ is the deformation, and Φ is a signal representation (e.g., Fourier transform or wavelet transform). This condition indicates that the distance between the representations $\Phi(x)$ and $\Phi(x_\tau)$ should be of the order of the deformation and that time-warping deformations are locally linearized by $\Phi(x)$. In addition, this condition suggests global translation invariance if $\tau(t) = c$, then the derivative of τ becomes $\tau'(t) = 0$, which is a pure translation without deformations.

The Fourier transform and the wavelet transform are signal representations that do not provide invariance to translations and stability to small time-warping deformations. The magnitude of the Fourier transform is invariant to

translations. The above can be proven by considering a signal $x(t) \in L^2(\mathbb{R})$ of finite energy $\int_{-\infty}^{\infty} |x(t)|^2 dt < +\infty$, whose Fourier transform can be obtained as shown in (2):

$$\hat{x}(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt \quad (2)$$

On the other hand, if $x_c(t) = x(t - c)$ is the translated version of $x(t)$, the Fourier transform of $x_c(t)$ can be obtained by considering the time-shifting property of the Fourier transform which leads to the expression shown in (3):

$$\hat{x}_c(\omega) = e^{-ic\omega} \hat{x}(\omega) \quad (3)$$

Based on (3), it can be appreciated that when a translation is applied to a signal, it introduces a phase shift to the Fourier transform of $x(t)$ through the term $e^{-ic\omega}$. However, this phase shift in the Fourier domain does not change the magnitude of the Fourier transform. This implies that the Fourier transform magnitudes of $x(t)$ and $x_c(t)$ are the same, as shown in (4):

$$|\hat{x}(\omega)| = |\hat{x}_c(\omega)| \quad (4)$$

However, if a signal is deformed, the magnitude of the Fourier transform changes and becomes unstable. The simplest form of deformation is a translation that depends on time $\tau(t) = \epsilon t$, with $0 < \epsilon \ll 1$. Consequently, $x(t)$ deformed can be expressed as shown in (5):

$$x_\tau(t) = x(t - \tau(t)) = x((1 - \epsilon)t) \quad (5)$$

Hence, considering the scaling property of the Fourier transform, the frequency representation of $x_\tau(t)$ is presented in (6):

$$\hat{x}_\tau(\omega) = \frac{1}{1 - \epsilon} \hat{x}\left(\frac{\omega}{1 - \epsilon}\right) \quad (6)$$

This deformation can be understood as a dilation that shifts the signal's frequency components. Since the deformation ϵ is small, the higher frequencies of the signal will become higher compared to the lower frequencies due to the factor $\frac{1}{1-\epsilon}$.

Based on the relation provided in (6), it is possible to observe that the Fourier transform does not provide stability to deformations by considering small deformations. On the other hand, the wavelet transform is defined in (7):

$$Wx(u, s) = \langle x(t), \psi_{u,s}(t) \rangle = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt \quad (7)$$

where Wx is the wavelet transform of $x(t)$, $*$ is the complex conjugate, and $\psi_{u,s}(t) \in L^2(\mathbb{R})$ represents the mother wavelet used to perform the wavelet transform for a particular scale s , and translation u . The scale parameter allows controlling the frequency localization, while the translation parameter controls the time domain localization [56]. A wavelet is a localized waveform that is stable to deformations contrary to the sinusoids of the Fourier transform. Nevertheless, since the wavelet transform depends on time, this transform is not translation invariant; on the contrary, it is translation covariant [45].

The wavelet transform's translation covariant nature can be understood as a convolution product between a signal and a wavelet at a specific frequency scale, as shown in (8):

$$Wf(u, s) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^*(\frac{t-u}{s}) dt = f \star \bar{\psi}_s(u) \quad (8)$$

with

$$\bar{\psi}_s(t) = \frac{1}{\sqrt{s}} \psi^*(\frac{-t}{s}) \quad (9)$$

where the \star denotes a convolution product. Since the wavelet transform relies on a convolution between the signal of interest and the mother wavelet, it is inherently time-dependent. Thus, it is possible to observe that neither the Fourier transform nor the wavelet transform provides invariance to translations and stability to small deformations. Therefore, to construct an invariant representation of a signal stable to small time-warping deformations, the WST employs convolution through wavelets, non-linearities applied through modulus operations, and averaging with scaling functions. Taking the above into account, the WST is constructed by defining a scaling function $\phi(t)$ and a family of dilated wavelets $\psi_\lambda(t)$. The key results of the WST can be named as zero, first, and second-order scattering coefficients, whose computation process is described below [45].

The scaling function $\phi(t)$ is a low-pass filter that computes the invariant representation by averaging the scattering coefficients at a particular scale defined by 2^J . The zero-order scattering coefficients $S_0x(t)$ are computed by taking the convolution of the input signal $x(t)$ and the scaling function $\phi(t)$ [57]. This process is expressed in (10):

$$S_0x(t) = x(t) \star \phi(t) \quad (10)$$

This process produces local translation invariant descriptors but removes the high-frequency content of $x(t)$. Hence, the critical challenge is constructing an invariant without removing the signal's high-frequency information since discriminative information could be present in its higher frequencies.

The wavelets $\psi_\lambda(t)$ work as band-pass filters ($\hat{\psi}(0) = 0$) that cover the rest of the frequency spectrum that was not covered by the scaling function $\phi(t)$ [45]. These wavelets are considered complex wavelets with a quadrature phase and approximately analytical such as $\hat{\psi}(\omega) \approx 0$ for $\omega < 0$. Analytical wavelets are preferred to represent the frequency transients of oscillatory signals, contrary to real wavelets, which are better suited for representing sharp signal transitions [56], [58]. On the other hand, $\psi_\lambda(t)$ are wavelets dilated at different frequencies. For any $\lambda > 0$, a dilated wavelet of center frequency λ is obtained as shown in their time and frequency domain representation presented in (11) and (12):

$$\psi_\lambda(t) = \lambda \psi(\lambda t) \quad (11)$$

$$\hat{\psi}_\lambda(\omega) = \hat{\psi}\left(\frac{\omega}{\lambda}\right) \quad (12)$$

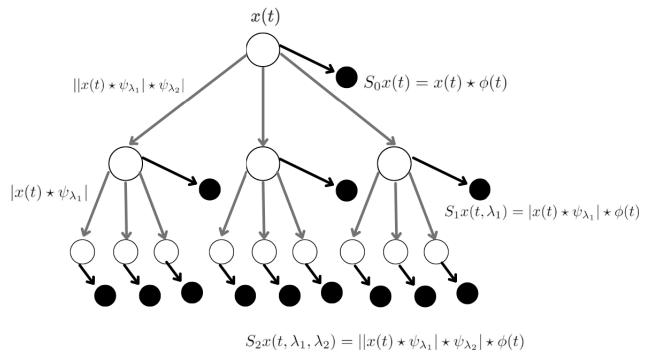


FIGURE 3. WST schematic representation. This image was adapted from [57].

These wavelets are used to recover the high-frequency information discarded from the signal when computing the zero-order scattering coefficients $S_0x(t)$; the frequency resolution of this family of wavelets is controlled by a factor Q that defines the number of wavelets per octave within a frequency octave or in other words the number of filters per octave. The center frequency of the wavelets is normalized to 1. Considering that the Q factor represents the number of wavelets per octave, lambda is equivalent to $\lambda = 2^{(k/Q)}$ for $k \in \mathbb{Z}$. The bandwidth of $\hat{\psi}$ is in the order of Q^{-1} . Furthermore, the support of $\hat{\psi}_\lambda(\omega)$ is centered in λ with a bandwidth given by λ/Q , and the energy of the wavelets is concentrated around 0 in the interval of size $2\pi Q/\lambda$ [45], [57].

The first-order coefficients are obtained by first making the convolution with the family of wavelets labeled as ψ_{λ_1} dilated at different octaves controlled by Q_1 with the input signal and then taking the modulus of the convolution result. Computing the modulus of analytical wavelet coefficients can be interpreted as a sub-band Hilbert envelope demodulation [45]. Finally, the modulus results are convolved again with the scaling function $\phi(t)$. This process is described in the expression shown in (13):

$$S_1x(t, \lambda_1) = |x(t) \star \psi_{\lambda_1}| \star \phi(t) \quad (13)$$

To obtain the second-order coefficients, the coefficients obtained by convolving the input signal with the first family of wavelets ψ_{λ_1} and subsequently computing the modulus are convolved again with a second family of wavelets ψ_{λ_2} . There is only one wavelet per octave for this second layer, which implies that $Q_2 = 1$. Consequently, the modulus is applied to the convolution result with this second family of wavelets, and the result of applying the modulus is convolved again with the scaling function. These operations are described in the expression shown in (14):

$$S_2x(t, \lambda_1, \lambda_2) = ||x(t) \star \psi_{\lambda_1} \star \psi_{\lambda_2}| \star \phi(t) \quad (14)$$

The WST can be extended to further layers; nevertheless, the signal's energy dissipates with more layers, becoming zero. Based on the works of Bruna and Mallat [59],

and Andén and Mallat [45], for most applications, two layers are sufficient since the energy of the signals starts dissipating with more layers [57]. This process progressively translates higher frequencies to lower frequencies through convolutions with wavelets, non-linearities applied via modulus operations, and convolution with scaling functions, which, consequently, achieves stability to small time-warping deformations. Fig. 3 illustrates the scattering transform operations in a graphic representation [45], [59].

The WST is similar to CNNs in that both use a cascade of convolutions to process the input signal. However, there are some differences between the two. In WST, wavelets are used as fixed filters for convolution, whereas CNNs learn filters from the data. Additionally, the number of layers and the architecture of WST is predetermined, whereas in CNNs, the architecture is determined empirically. Finally, the modulus operation in WST is similar to the non-linearities used in CNNs. At the same time, the convolution with the scaling function in each scattering layer in WST is akin to the pooling operation of CNNs [57]. Please refer to the works of Andén and Mallat [45], Bruna and Mallat [59], and Mallat [55] for further details related to the WST.

C. MACHINE LEARNING TECHNIQUES

This section presents an overview of the machine learning techniques used to classify water leakages by employing the WST as a feature extraction technique on vibration, dynamic pressure, and hydrophone data. Classical machine learning classifiers were selected to maintain the final model's simplicity and avoid using complex classification techniques such as random forest, multilayer neural networks, XGboost, or CNNs [60]. Although the techniques above could provide higher performance, they require a large sample size of training data and high computational power. Besides, the complexity of these techniques (e.g., the number of trees of a random forest or the number of layers and neurons of deep neural networks) hinders the resultant model's interpretability [61]. Considering the above, the computed WST features were used to train a multiclass SVM, a KNN, and a DT.

SVMs were selected since they generate linear decision boundaries to discriminate between classes [57]. The KNN was selected due to the algorithm's simplicity in making a prediction [62]. DTs are a type of algorithm that classify data based on a set of rules. These rules are designed in a way that humans can easily understand them. Therefore, DTs facilitate interpreting the results obtained from the generated model [63]. In addition, SVMs, DTs, and KNNs are frequently evaluated in related research when performing water leak detection [19], [23], [25], [26], [44], [47], [64].

The machine learning techniques used in this study were trained using the MATLAB 2023b implementation. The machine learning techniques and the WST computation were performed on a computer with an Intel(R) Core(TM) i7-7700HQ CPU of 2.80GHz and 8.00 GB of RAM.

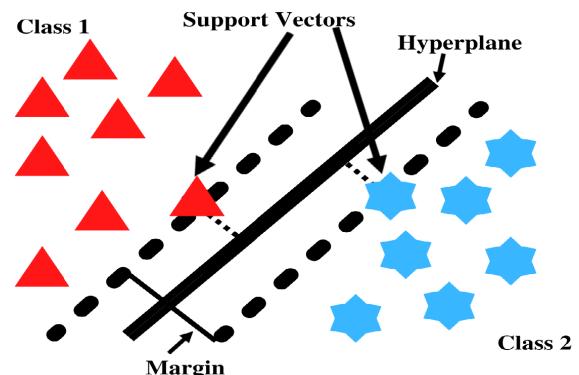


FIGURE 4. SVM schematic representation. This image was adapted from [57].

1) SUPPORT VECTOR MACHINE

The main objective of an SVM is to find a hyperplane that best separates two classes of data points. The best hyperplane is the one that has the widest margin between the two classes. The margin is the distance between two data points known as support vectors. Fig. 4 illustrates the hyperplane, support vectors, and margin generated by an SVM by considering a linear decision boundary. An SVM is defined for binary classification problems, so the problem is decomposed using multiple binary SVM classifiers for multiclass classification. For non-linearly separable data, a kernel projects the data to a higher dimensional space in which the compared classes can be linearly separated. A typical kernel used for this purpose is the RBF [65].

2) K-NEAREST NEIGHBORS

The KNN classification algorithm compares a test set to similar training sets with n features. When a new test data point is going to be classified, the KNN searches for the k training points that are nearer to the unknown test set in a n -dimensional feature space. A distance metric (e.g., Euclidean distance) is employed to determine the proximity of the test data point to the training set. A majority voting mechanism is then used to decide the most prevalent class among the test set. The KNN algorithm is a lazy classifier because it memorizes the training dataset rather than producing a model from the training set [66].

3) DECISION TREE

DTs, also known as classification trees, are a type of machine-learning technique used to predict qualitative responses. These techniques are classified within the broader class of tree-based methods, such as regression trees, random forests, and gradient-boosted trees. A DT classifies new data points based on the most frequent class in a partitioned region. Hence, the crucial task in generating DTs is constructing the best partition. DTs partition data based on impurity. If a subset contains only one class, it is pure; otherwise, it is impure. The quality of a split is determined by its purity. Two

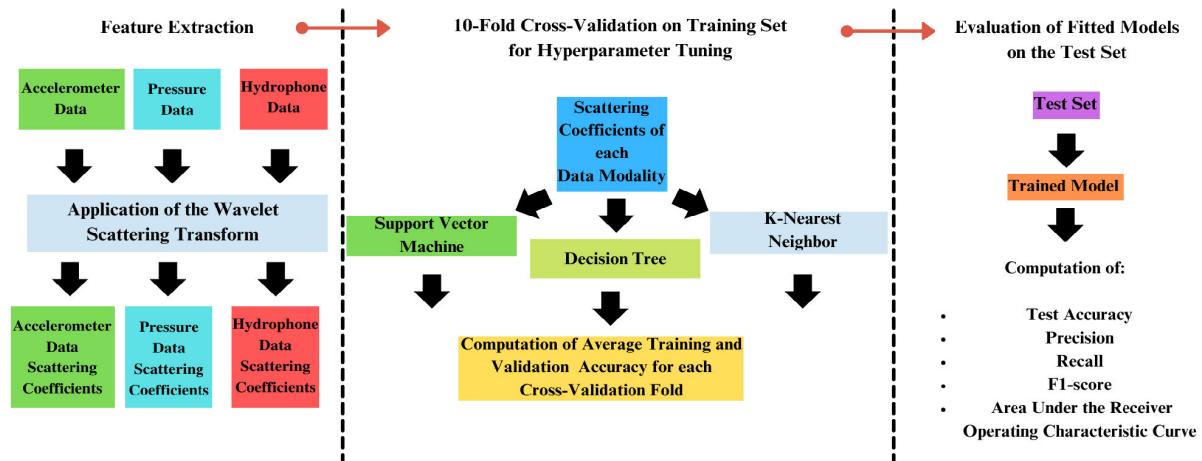


FIGURE 5. This is an outline of the steps considered to compare the performance of each data modality for water leak classification. First, the WST is applied to each data modality to obtain features that allow machine learning techniques to be trained. Consequently, the hyperparameters of an SVM, DT, and KNN are set by applying a 10-fold cross-validation on a training set of each data modality WST. Finally, the trained models are evaluated on a test set containing each WST data modality.

metrics are frequently used to determine the best partition: the Information Gain and the Gini Impurity Index. Once the classification tree has been constructed, a test data point is compared with a threshold value in case the input data is numeric or with a range of possibilities for the case of nominal data points [67].

The performance of these techniques is evaluated when trained with features derived from applying the WST to hydrophone, accelerometer/vibration, and dynamic pressure data by employing 10-fold cross-validation. Consequently, the generated models were evaluated on a test set to evaluate their performance. Fig. 5 presents an outline of the steps considered to evaluate the performance of each data modality for water leak classification [57].

D. PREPROCESSING

A total of 8 measurements per leak type were recorded for 30 seconds for both accelerometer and dynamic pressure data. These eight measurements represent the response of the two sensors considered for each data modality within the water networks while a water demand was generated via the service line (details provided in Section III and Table 2). This ensures an equal distribution of measurements across leak types: 8 for the circumferential crack leak, 8 for the longitudinal crack leak, 8 for the orifice leak, 8 for the gasket leak, and 8 for no-leak scenarios. Considering the above, the measurements of each data modality were preprocessed before applying the WST to the data to train the selected machine learning classifiers. Fig. 6 shows an overview of the preprocessing steps applied to each measurement.

For each water network topology, 40 measurements were recorded for both accelerometer and dynamic pressure data, considering the two accelerometer and dynamic pressure sensors. Each 30-second recorded measurement was then segmented using non-overlapping windows of 1 second

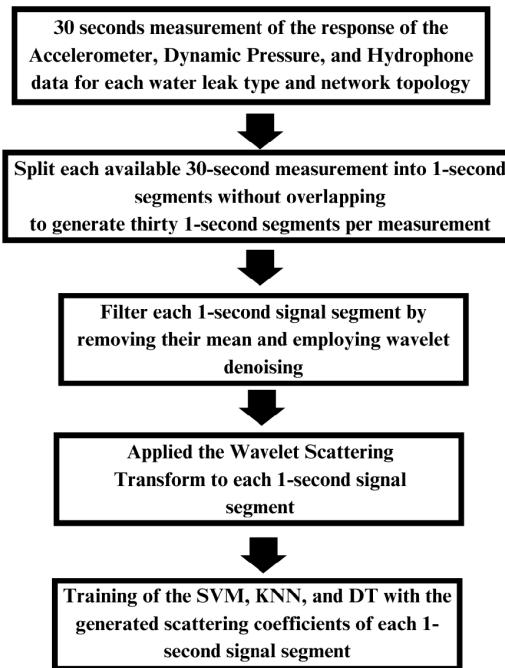


FIGURE 6. Preprocessing steps applied to each of the 30-second measurements of the accelerometer, dynamic pressure, and hydrophone data available in the dataset.

to augment the dataset's sample size. This segmentation approach generated 30 signal segments per measurement, resulting in 240 signal segments per leak type and sensor type and 1200 signal segments for each water network topology and sensor type.

For the case of the hydrophone data, a total of 8 measurements (4 measurements related to H1 and 4 to H2) were recorded per leak type that considered background noise while being recorded. The dataset contained four other

TABLE 3. Number of 1-second signal segments after segmenting each 30-second measurement.

Signal segments	1-Second Signal segments
Signal segments per 30-second measurement	30
Signal segments per leak type (considering the two sensors available for each data modality and the four background water demands)	240
Signal segments per data modality	1200
Signal segments per water network topology (considering the three data modalities)	3600

measurements (2 measurements related to H1 and 2 to H2) that did not contain background noise while being recorded (see Table 2). Nevertheless, these four measurements were discarded to maintain an equal sample size for each data modality. Similar to the vibration and dynamic pressure data, each recorded measurement of the hydrophone data was segmented by employing non-overlapping windows of 1 second. This led to generating 30 signal segments per measurement, leading to 240 signal segments per leak type and 1200 signal segments for each water network topology. Table 3 summarizes the number of signal segments generated for each leak type, sensor type, and water network after segmenting them.

Before applying the WST to the generated signal segments of the accelerometer, hydrophone, and dynamic pressure data, a denoising stage was performed on each signal to remove the mean and high-frequency noise of the signal segments. For removing the mean of each 1-second signal segment, the expression shown in (15) was considered:

$$X_{ij} = Y_{ij} - \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij} \quad (15)$$

where X_{ij} is the j th data sample of the i th signal segment of the dataset of size N after removing the mean; Y_{ij} is the j th sample of the original i th raw signal without removing the mean. Finally, n_i is the size of the i th signal segment [52].

A wavelet denoising approach was utilized to remove outliers and high-frequency noise from the generated signal segments. Wavelet denoising was chosen over other filtering techniques, such as infinite and finite impulse response or Fourier-based filters, since it is unnecessary to know the frequency band or bands in which noise is present in the data a priori. Wavelet denoising consists of applying the discrete wavelet transform on the signal to be filtered. The obtained wavelet coefficients are thresholded based on a noise estimation technique and thresholding rule. Once the wavelet coefficients are thresholded, the inverse discrete wavelet transform is applied to the thresholded wavelet coefficients to obtain the denoised signal [68], [69]. Wavelet denoising has been used effectively in several applications since the wavelet transform produces a sparse representation or very few large coefficients that may be uncorrelated with noise [70]. This study utilized a Symlet of order 3 to compute the discrete wavelet transform. The noise estimation

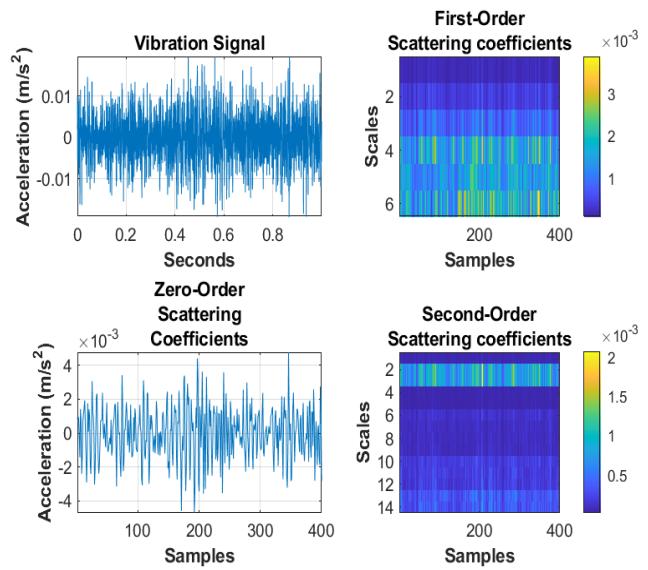


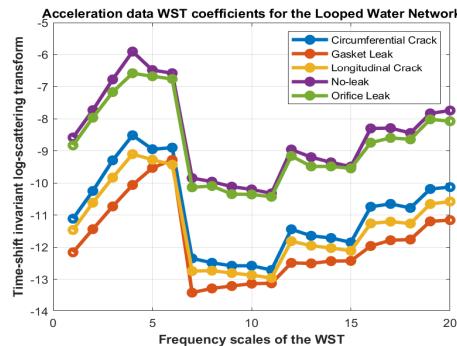
FIGURE 7. WST coefficients of vibration signals associated with a water leak. The top of the Figure shows the original vibration signal (left) and the first-order scattering coefficients (right). The bottom shows the zero-order scattering coefficients (left) and the second-order scattering coefficients (right).

technique was a universal threshold, and the decomposition level was set to 4. The thresholding rule adopted was a soft thresholding [52].

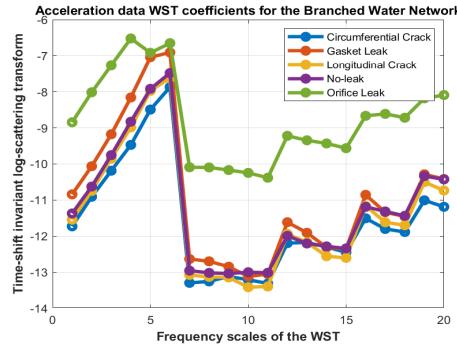
E. FEATURE EXTRACTION OF VIBRATION DATA

This work computed the WST with the implementation of MATLAB 2023b. Two key parameters are set to compute the WST: the number of wavelets per octave (Q) to compute the first- and second-order scattering coefficients and the invariance scale. It is important to note that the number of wavelets per octave allows for the control of the frequency resolution when computing the wavelet transform in each scattering layer. In contrast, the invariance scale controls the size of the time-averaging window of the scaling function. For the first and second-order scattering coefficients, the number of wavelets per octave was set as $Q_1 = 1$ and $Q_2 = 1$, respectively. For the case of the invariance scale, a value of 0.01 seconds was set. The selection of one wavelet per octave and the invariance scale of 0.01 seconds were chosen to produce a low number of input features; even though a higher number of wavelets per octave can produce a better frequency resolution, the number of scattering coefficients increases as well as the number of input features to train the machine learning classifiers, which can lead to overfitting. Fig. 7 shows the obtained zero, first, and second-order scattering coefficients after applying the scattering transform to one of the signal segments of the accelerometer data.

Following the recommendation and guidelines of the Kymatio library, the zero-order scattering coefficients were not considered as input features [71]. The obtained coefficients were transformed via the natural logarithm, generating the log-scattering transform. Consequently, the coefficients



(a)



(b)

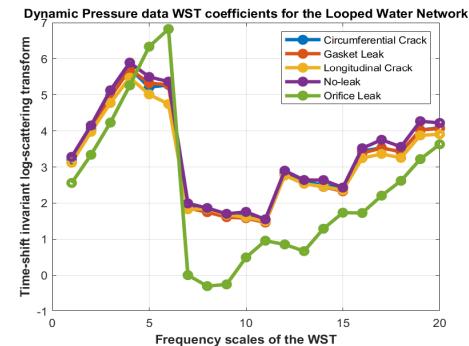
FIGURE 8. Wavelet scattering coefficients obtained from the accelerometer data. a) Scattering coefficients of the looped water network. b) Scattering coefficients of the branched water network.

were averaged along the time dimension. This process produces an utterly invariant representation since the scattering coefficients along the time dimension were aggregated. This process is equivalent to the one presented in [57].

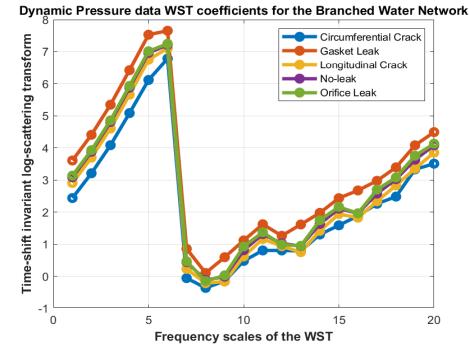
Fig. 8 shows the scattering coefficients after applying the natural logarithm and averaging them over time. This representation receives the name of the time-shift invariant log-scattering transform. A total of 20 coefficients were obtained following this procedure, which were then used as input to train the selected machine learning classifiers. This process was repeated for each 1-second signal segment of the dataset.

F. FEATURE EXTRACTION OF DYNAMIC PRESSURE DATA

The parameters used to compute the scattering transform on the dynamic pressure data were the same as those used for the accelerometer data, with the number of wavelets per octave for the first and second-order scattering coefficients set to $Q_1 = 1$ and $Q_2 = 1$, respectively and the invariance scale was set to 0.01 seconds. Furthermore, similar to the scattering coefficients obtained from the accelerometer data, the zero-order scattering coefficients were not considered as input features, so they were removed following the Kymatio library recommendation. Subsequently, the first- and second-order scattering coefficients were log-transformed and averaged over time to produce the time-shift invariant log-scattering transform. This process is equivalent to the one performed on the accelerometer/vibration data. An example



(a)



(b)

FIGURE 9. Wavelet scattering coefficients obtained from the dynamic pressure data. a) Scattering coefficients of the looped water network. b) Scattering coefficients of the branched water network.

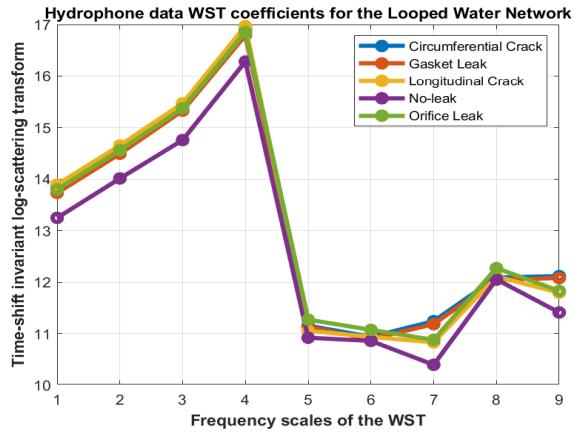
of the obtained scattering coefficients of the dynamic pressure data associated with each leak type is presented in Fig. 9 for the looped and branched water networks. Similar to the accelerometer data, the number of coefficients obtained was equal to 20, which were used to train the SVM, KNN, and DT.

G. FEATURE EXTRACTION OF HYDROPHONE DATA

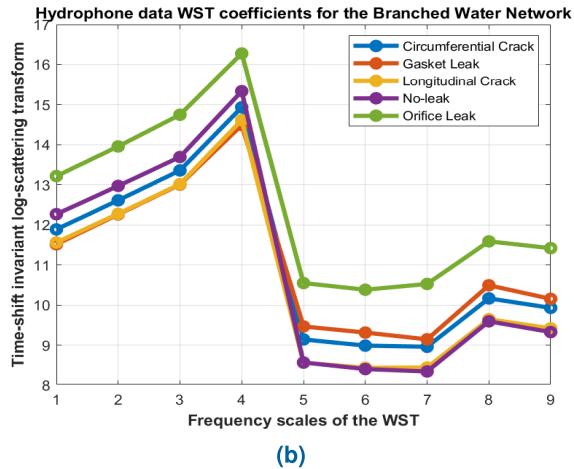
Similar to the other data modalities employed in this study, to apply the WST to the hydrophone data, the number of wavelets per octave for the first and second-order scattering coefficients was set as $Q_1 = 1$ and $Q_2 = 1$, respectively, and the invariance scale was set to 0.01 seconds. Furthermore, the zero-order scattering coefficients were discarded for the analysis following the Kymatio library suggestion, similar to the accelerometer and dynamic pressure data zero-order scattering coefficients. The scattering coefficients were also log-transformed and averaged over time to obtain the time-shift invariant log-scattering transform. Fig. 10 shows the obtained scattering coefficients of the hydrophone data for each water leak type for the looped and branched water networks. By following the procedure above, a total of 9 coefficients were obtained. These coefficients were used to train the SVM, KNN, and DT.

H. MULTIMODAL WATER LEAKAGE CLASSIFICATION

This study utilized the trained machine learning classifiers that performed best on the training set for each data modality



(a)



(b)

FIGURE 10. Wavelet scattering coefficients obtained from the hydrophone data. a) Scattering coefficients of the looped water network. b) Scattering coefficients of the branched water network.

and water network to create an ensemble model incorporating all sensor types as input for categorizing water leakages. The multimodal model to achieve this was implemented using LF. LF addresses different data modalities by employing a distinct model for each data type. Subsequently, a voting mechanism, weighing or averaging, combines the results of each data modality model. This study used a majority voting technique (hard voting) to aggregate the predictions of the models trained for each data modality [72], [73], [74]. LF was chosen over other multimodal techniques due to its simplicity and ability to leverage each data modality's previously trained models. This approach is illustrated in Fig. 11. The hydrophone, accelerometer, and dynamic pressure sensor measurements were assumed to be sampled synchronously for each leak type and scenario based on the information provided in [53] to apply the LF approach and merge the prediction of the classifiers trained for each data modality.

IV. RESULTS

This work divided the dataset into two main sets: training and testing sets. 80% of the dataset was used for training,

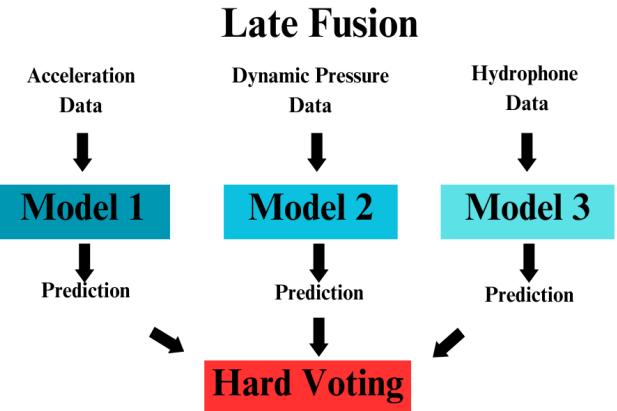


FIGURE 11. LF schematic representation.

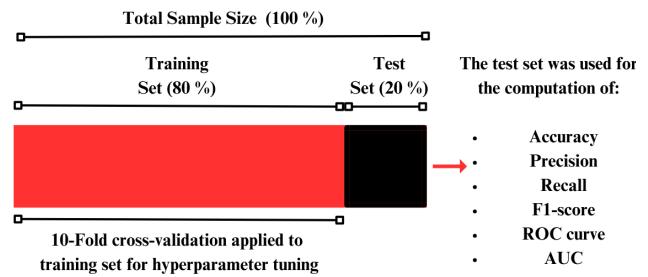


FIGURE 12. Validation process used to select the hyperparameters of the machine learning techniques used in this study and evaluate their performance. ROC refers to the receiver operating characteristic curve. AUC refers to the area under the ROC curve.

while 20% was used for testing. This partition of the dataset was randomly generated. The training set was used for hyperparameter tuning the selected machine learning techniques. The hyperparameter tuning was set by applying 10-fold cross-validation on the training set. 10-fold cross-validation consists of randomly splitting the dataset into ten equal-sized parts, where nine parts are used for training the model and 1 for validating the trained models. This process is repeated ten times, and the training and validation sets are interchanged between the ten parts. The performance obtained in each fold is saved and finally averaged to assess the model's generalization performance. In this study, 10-fold cross-validation was selected since it is commonly used for assessing the generalization of a machine learning model [75]. Fig. 12 illustrates the splitting process of the dataset conducted for evaluating the performance of each model presented in this study.

The performance of the fitted models was measured using the Accuracy as shown in (16), the Recall/Sensitivity as shown in (17), the Precision as shown in (18), and the F1-score as shown in (19). The equations below use four metrics: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The F1-score, Precision, and recall were reported only for the testing set. Accuracy was reported for both the training and validation sets to evaluate the overfitting of each machine technique for

different hyperparameter values [52], [57].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (16)$$

$$\text{Recall/Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (17)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (18)$$

$$\text{F1-score} = \frac{2(\text{Precision})(\text{Recall})}{\text{Precision} + \text{Recall}} \quad (19)$$

The recall, precision, and F1-score were computed for each class or leak type using the testing set to assess the trained classifiers' performance on each water leak type. The accuracy of the trained classifiers on the testing set was also computed. The average training and validation accuracy with the corresponding standard deviation was computed using the 10-fold cross-validation on the training set.

The Area Under the Receiver Operating Characteristic Curve (AUC) was also considered to evaluate the performance of the proposed method for water leakage classification. The Receiver Operating Characteristic (ROC) Curve illustrates how a classification model balances its sensitivity and specificity for any classification threshold. A perfect classifier is represented by an AUC value of 1, while a value of 0.5 represents a random classifier [52], [76].

A. WATER LEAK CLASSIFICATION RESULTS BASED ON VIBRATION DATA

An RBF kernel was utilized to train the SVM using the accelerometer data WST coefficients presented in Fig. 8. The SVM trained with RBF kernels possesses two hyperparameters, C and gamma. The gamma hyperparameter controls the degree of curvature of the decision boundary generated by the SVM. On the other hand, the C hyperparameter controls the model's error degree [77]. The value of the gamma hyperparameter was fixed to 1 for the models trained for the looped and branched water networks. The candidate values of the regularization hyperparameter C were $C = [0.001, 0.1, 1, 10, 50, 100, 250, 500, 750, 1000]$. The regularization hyperparameter C values were set exponentially as suggested in [57]. A z-score standardization was applied to the scattering coefficients of the accelerometer data. Standardizing data when variables have different ranges is essential to avoid giving a higher weight to variables with broader ranges when using distance-based techniques [78].

Fig. 13 shows the multiclass SVM average training and validation accuracies when trained with the accelerometer data scattering coefficients by changing the regularization hyperparameter C value for the looped water network and branched water network models. The vertical dashed line presented in Figs. 13(a) and 13(d) for each validation curve denotes the hyperparameter that provides the best average validation accuracy when performing 10-fold cross-validation on the training set [77]. This C value was used to train and evaluate the final model with the test set.

For the case of the KNN, the candidate set of neighbors (n) were $n = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$. The Euclidian distance was the metric for evaluating the nearest neighbors. A z-score standardization was applied to the input features. Figs. 13(b) and 13(e) show the average training and validation accuracy of the KNN by being trained with the WST of the accelerometer data presented in Fig. 8 by applying 10-fold cross-validation to the training set to determine the adequate number of neighbors for the looped and branched water network models. The vertical dashed black line denotes the neighbor value that produces the best average validation accuracy. This neighbor value was used to train and evaluate the final model with the test set.

For the DT trained with the scattering coefficients of accelerometer data, the minimum leaf size of the tree was controlled by evaluating the following candidate set leaf size = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50]. The minimum leaf size is an essential hyperparameter utilized in DT algorithms, which refers to the minimum number of data points required to form a leaf node during the tree-building process. By setting a minimum leaf size, the complexity of the tree can be regulated to avoid overfitting, which is crucial in ensuring the accuracy and reliability of the DT model [79]. Therefore, carefully selecting an appropriate minimum leaf size can significantly enhance the effectiveness of the DT algorithm, particularly in data-rich environments. Contrary to the SVM and KNN, this algorithm does not require preprocessing the input features. Thus, a standardization process was not applied to the input features. The split criterion to construct the DT was the Gini impurity index. Figs. 13(c) and 13(f) show the DT's average training and validation accuracies by varying the tree's minimum leaf size and applying 10-fold cross-validation to the training set. The depth of the tree was not controlled in this study since MATLAB 2023b does not allow it to be controlled. The vertical dashed line of Figs. 13(c) and 13(f) show the minimum leaf size that produces the best average validation accuracy. This minimum leaf size value was used to train and evaluate the final model with the test set.

Table 4 displays the results of the multiclass SVM, DT, and KNN, for the test set. Table 4 also shows the hyperparameters that produced the best average validation accuracy, as indicated by the vertical dashed black lines in Fig. 13. The corresponding average training and test accuracy are also included in Table 4 and the individual class performance regarding recall, precision, F1-score, and AUC. Fig. 14 shows the ROC curves with the corresponding AUC of each machine learning technique obtained by testing the fitted models with the test set for the looped and branched water network models.

B. WATER LEAK CLASSIFICATION RESULTS BASED ON DYNAMIC PRESSURE DATA

For the SVM trained with the dynamic pressure data scattering coefficients (see Fig. 9), the same parameterization

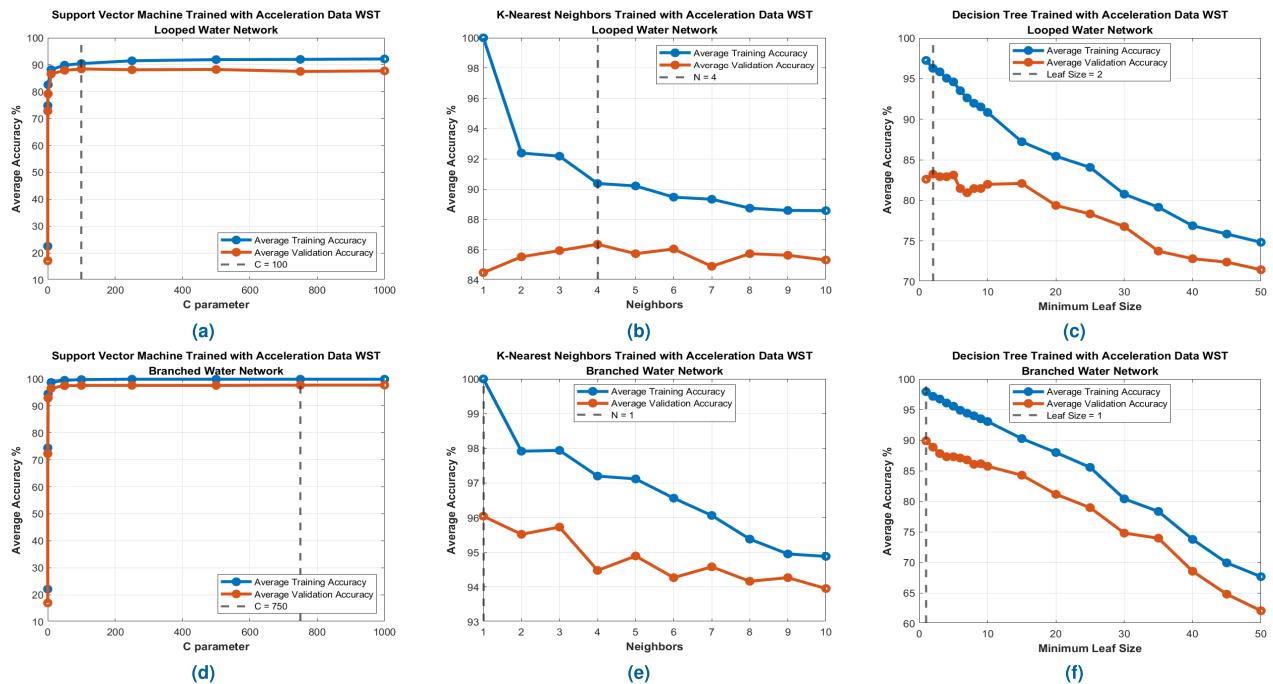


FIGURE 13. Average accuracies of training (blue) and validation (red) for different hyperparameters of the SVM, KNN, and DT trained using vibration signal features derived from the WST for water leakage classification. The vertical dashed line indicates the hyperparameter value that results in the best trade-off between the average accuracies of the training and validation sets. The top graphs display the models trained for the looped water network, and the bottom graphs display the models trained for the branched water network.

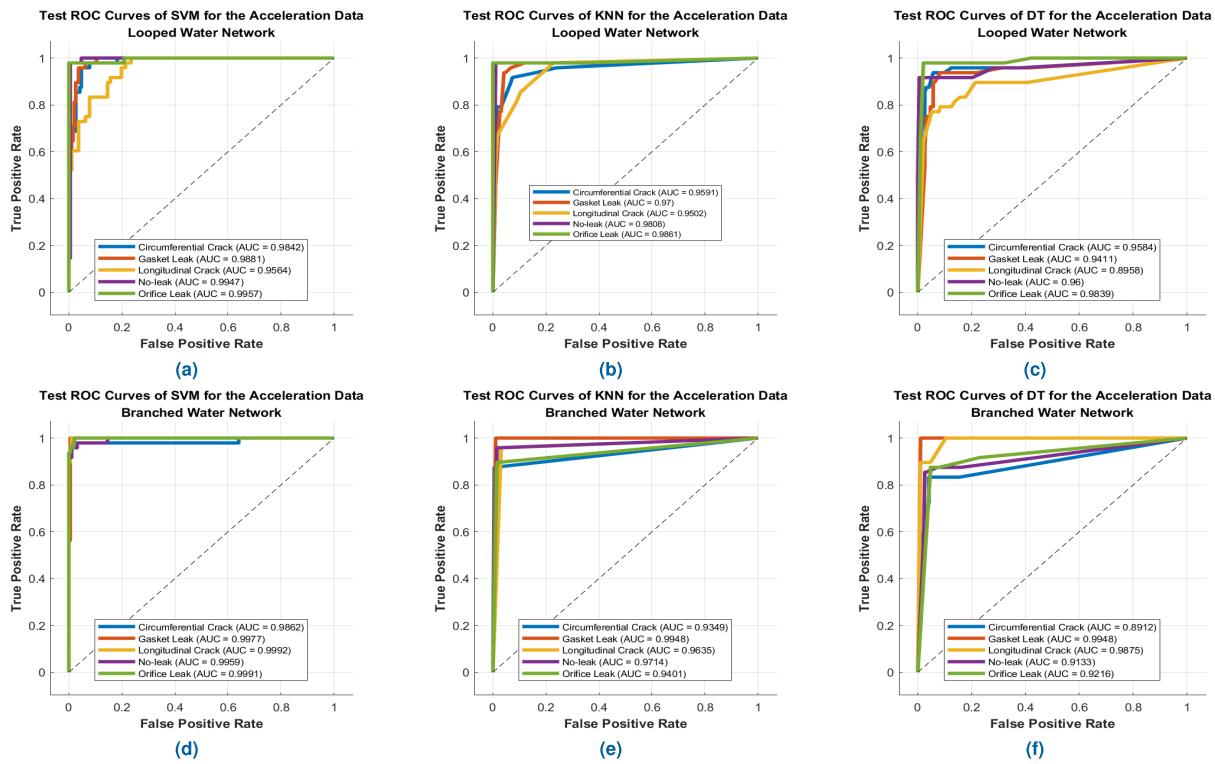


FIGURE 14. Test ROCs and AUCs of each machine learning technique for the looped (top) and branched (bottom) water networks trained models.

was used as the SVM trained with the accelerometer data scattering coefficients. This was done by setting gamma equal to 1 and the same candidate values of the regularization

hyperparameter C were used $C = [0.001, 0.1, 1, 10, 50, 100, 250, 500, 750, 1000]$. A z-score standardization was applied to the input features. Figs. 15(a) and 15(d) show

TABLE 4. Results of testing the SVM, KNN, and DT trained with the scattering coefficients of accelerometer data for water leakage classification for the looped and branched water network models.

Water Network	Machine Learning Technique	Parameters	Accelerometer Data							
			Average Training Accuracy %	Average Validation Accuracy %	Testing Accuracy %	Water Leak Type %	Test Recall %	Test Precision %	Test F1-score %	Test AUC
Looped	SVM	Kernel: RBF Gamma: 1 $C = 100$	90.4282 \pm 0.3538	88.4375 \pm 2.9646	89.5833	Circumferential Crack	89.5833	82.6923	86.00	0.9842
						Gasket Leak	89.5833	89.5833	89.5833	0.9881
						Longitudinal Crack	72.9167	77.7778	75.2688	0.9564
						No-leak	97.9167	97.9167	97.9167	0.9947
						Orifice Leak	97.9167	100	98.9474	0.9957
		Kernel: RBF Gamma: 1 $C = 750$	99.8958 \pm 0.0366	97.7083 \pm 1.5372	96.6667	Circumferential Crack	93.75	97.8261	95.7447	0.9862
						Gasket Leak	100	97.9592	98.9691	0.9977
						Longitudinal Crack	100	96	97.9592	0.9992
						No-leak	95.8333	93.8776	94.8454	0.9959
						Orifice Leak	93.75	97.8261	95.7447	0.9991
Branched	KNN	Neighbors: 4	90.3704 \pm 0.6046	86.3542 \pm 2.4281	89.1667	Circumferential Crack	89.5833	76.7857	82.6923	0.9591
						Gasket Leak	93.75	84.9057	89.1089	0.97
						Longitudinal Crack	66.6667	91.4286	77.1084	0.9502
						No-leak	97.9167	95.9184	96.9072	0.9808
						Orifice Leak	97.9167	100	98.9474	0.9861
		Neighbors: 1	100 \pm 0	96.1458 \pm 2.0276	93.75	Circumferential Crack	87.5	97.6744	92.3077	0.9349
						Gasket Leak	100	96	97.9592	0.9948
						Longitudinal Crack	95.8333	88.4615	92.000	0.9635
						No-leak	95.8333	93.8776	94.8454	0.9714
						Orifice Leak	89.5833	93.4783	91.4894	0.9401
Looped	DT	Leaf Size: 2	96.2616 \pm 0.8167	83.2292 \pm 3.6566	85.4167	Circumferential Crack	83.3333	88.8889	86.0215	0.9584
						Gasket Leak	75.0000	83.7209	79.1209	0.9411
						Longitudinal Crack	79.1667	67.8571	73.0769	0.8958
						No-leak	91.6667	97.7778	94.6237	0.96
						Orifice Leak	97.9167	92.1569	94.9495	0.9839
		Leaf Size: 1	97.8009 \pm 0.6316	90.4167 \pm 3.2498	88.75	Circumferential Crack	83.3333	83.3333	83.3333	0.8912
						Gasket Leak	100.0000	96.0000	97.9592	0.9948
						Longitudinal Crack	89.5833	95.5556	92.4731	0.9875
						No-leak	85.4167	87.234	86.3158	0.9133
						Orifice Leak	85.4167	82	83.6735	0.9216

the SVM's average training and validation accuracies while varying the regularization hyperparameter C for the looped and branched water network models, respectively. The black vertical dashed line denotes the hyperparameter value that produces the best validation accuracy. This hyperparameter value was used to train the final model and evaluate its performance with the test set.

For the KNN trained with the scattering coefficients of the dynamic pressure data, a z-score standardization was applied to the input features. An Euclidian distance was used to determine the nearest test data point. The candidate set of the nearest neighbors to label a test data point was $n = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$. Figs. 15(b) and 15(e) show the average training and validation accuracies for the KNN trained with the proposed candidate set of nearest neighbors for the looped and branched water network models, respectively. The dashed black line indicates the neighbor value that produces the best average validation accuracy. This hyperparameter value was used to evaluate the KNN with the test set.

Ultimately, the DT was trained by varying the minimum leaf size and using the scattering coefficients of the dynamic

pressure data as input. The split criterion to generate the DT was the Gini impurity index. The candidate set of minimum leaf size values was leaf size = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50]. No standardization was applied to the input features. Based on the above, the average training and validation accuracies obtained for the DT trained for different leaf size values are presented in Figs. 15(c), and 15(f) for the looped and branched water network models, respectively. The dashed black line shows the leaf size value that produces the best average validation accuracy. This minimum leaf size value was used to train the final DT and evaluate it with the test set.

Table 5 displays the SVM, KNN, and DT performance on the test set. The table also shows the hyperparameters that produced the best average validation accuracy, indicated by vertical dashed black lines in Fig. 15. The test accuracy and the individual class performance regarding recall, precision, F1-score, and AUC are also included in Table 5. The corresponding AUC values for each model can be found on the ROC curves depicted in Fig. 16 for the looped and branched water network models.

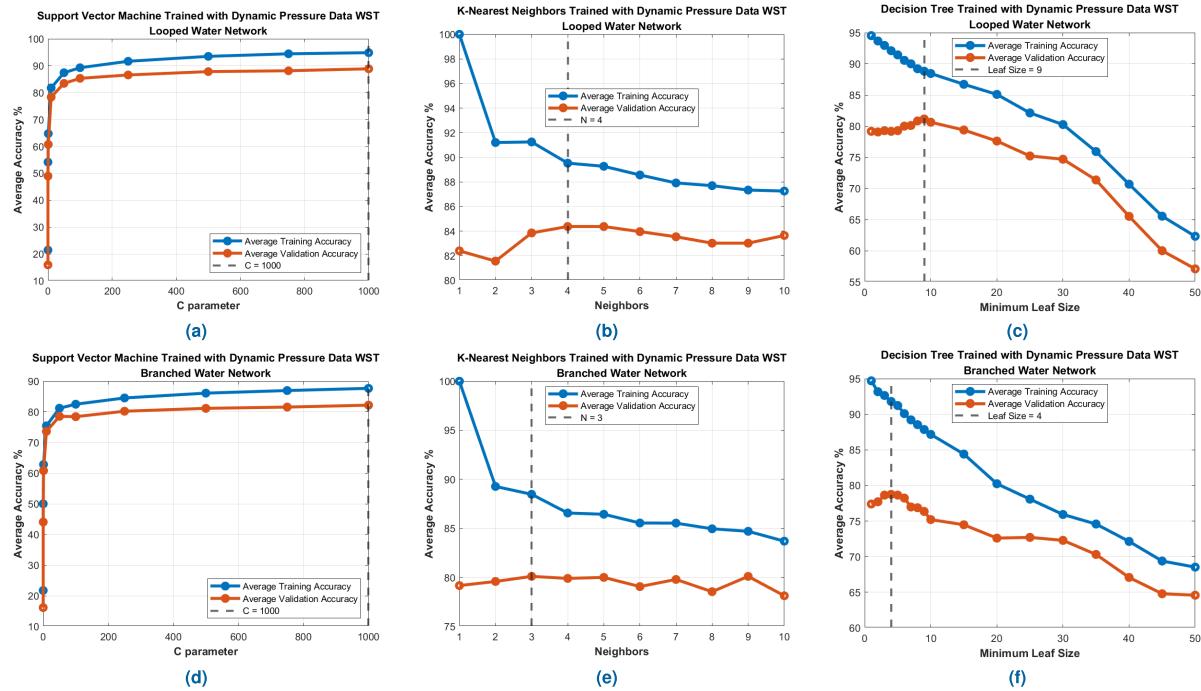


FIGURE 15. The graph displays training (blue) and validation (red) average accuracies for different hyperparameters used to train the SVM, KNN, and DT. These techniques were trained using dynamic pressure signal features derived from the WST to classify water leakage. The vertical dashed line marks the hyperparameter value that produces the best trade-off between the average accuracies of the training and validation sets. The top graphs show the results of models trained for the looped water network, while the bottom graphs show the results of models trained for the branched water network.

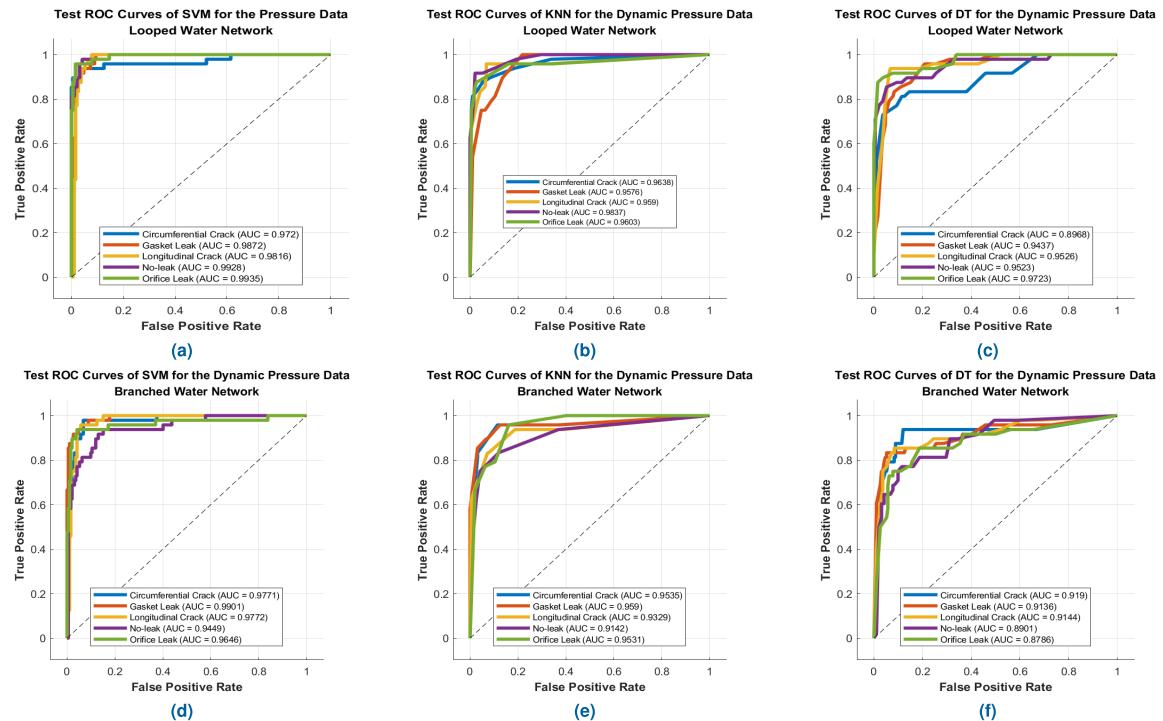


FIGURE 16. Test ROCs and AUCs of each machine learning technique trained for the looped (top) and branched (bottom) water networks.

C. WATER LEAK CLASSIFICATION RESULTS BASED ON HYDROPHONE DATA

The results of training the SVM, KNN, and DT with the scattering coefficients of the hydrophone data (see Fig. 10)

are presented in Fig. 17, and Table 6. The RBF was used to train the SVM, with gamma equal to 1, and the candidate set of values for the hyperparameter C were $C = [0.001, 0.1, 1, 10, 50, 100, 250, 500, 750, 1000]$, which

TABLE 5. Results of testing the SVM, KNN, and DT trained with the scattering coefficients of dynamic pressure data for water leakage classification for the looped and branched water network models.

Water Network	Machine Learning Technique	Parameters	Dynamic Pressure Data							
			Average Training Accuracy %	Average Validation Accuracy %	Testing Accuracy %	Water Leak Type %	Test Recall %	Test Precision %	Test F1-score %	Test AUC
Looped	SVM	Kernel: RBF Gamma: 1 C = 1000	94.8495 ± 0.4092	88.8542 ± 3.4391	90	Circumferential Crack	89.5833	95.5556	92.47	0.972
						Gasket Leak	93.75	83.3333	88.2353	0.9872
						Longitudinal Crack	89.5833	86	87.7551	0.9816
						No-leak	89.5833	89.5833	89.5833	0.9928
						Orifice Leak	87.5	97.6744	92.3077	0.9935
	Branched	Kernel: RBF Gamma: 1 C = 1000	87.6736 ± 0.8579	82.1875 ± 3.0449	84.5833	Circumferential Crack	93.75	77.5862	84.9057	0.9771
						Gasket Leak	87.5	93.3333	90.3226	0.9901
						Longitudinal Crack	75	85.7143	80	0.9772
						No-leak	75	81.8182	78.2609	0.9449
						Orifice Leak	91.6667	86.2745	88.8889	0.9646
Branched	KNN	Neighbors: 4	89.5255 ± 0.5096	84.375 ± 2.3032	81.6667	Circumferential Crack	87.5	80.7692	84	0.9638
						Gasket Leak	79.1667	69.0909	73.7864	0.9576
						Longitudinal Crack	85.4167	77.3585	81.1881	0.959
						No-leak	79.1667	92.6829	85.3933	0.9837
						Orifice Leak	77.0833	94.8718	85.0575	0.9603
	Looped	Neighbors: 3	88.4722 ± 0.6031	80.1042 ± 5.3443	79.1667	Circumferential Crack	83.3333	86.9565	85.1064	0.9535
						Gasket Leak	89.5833	76.7857	82.6923	0.959
						Longitudinal Crack	83.3333	74.0741	78.431	0.9329
						No-leak	62.5	85.7143	72.2892	0.9142
						Orifice Leak	77.0833	75.5102	76.2887	0.9531
Looped	DT	Leaf Size: 9	88.8194 ± 0.5754	81.1458 ± 3.16114	82.5	Circumferential Crack	72.9167	83.3333	77.7778	0.8968
						Gasket Leak	81.2500	72.2222	76.4706	0.9437
						Longitudinal Crack	93.7500	77.5862	84.9057	0.9526
						No-leak	77.0833	90.2439	83.1461	0.9523
						Orifice Leak	87.5000	93.3333	90.3226	0.9723
	Branched	Leaf Size: 4	91.8056 ± 0.4812	78.7500 ± 4.0848	75.4167	Circumferential Crack	79.1667	73.0769	76	0.919
						Gasket Leak	81.2500	79.5918	80.4124	0.9136
						Longitudinal Crack	83.3333	72.7273	77.6699	0.9144
						No-leak	60.4167	82.8571	69.8795	0.8901
						Orifice Leak	72.9167	71.4286	72.1649	0.8786

are the same as the ones used for the SVMs trained with vibration and dynamic pressure data. The above was done to test each data modality for water leakage classification in the same conditions using the same machine-learning techniques and perform a homogenous comparison of the results. The average training and validation accuracy of the SVM trained with the scattering coefficients of the hydrophone data can be appreciated in Figs. 17(a), and 17(d) for the looped and branched water network models, respectively. The vertical dashed black line shows the value of the regularization hyperparameter that produces the best average validation accuracy, which was used to train the final SVM and evaluate it with the test set.

For the KNN trained with the scattering coefficients of the hydrophone data, the candidate set of nearest neighbors was $n = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$. A z-score standardization was applied to the input features, and the distance metric to determine the nearest neighbors was an Euclidean distance. Figs. 17(b), and 17(e) show the average training and validation accuracies of the KNN while varying the number of nearest neighbors for the looped and branched

water network models, respectively. The vertical dashed black line shows the neighbor's value, which produces the best average validation accuracy. This value was used to train and evaluate the final KNN with the test set. Figs. 17(c), and 17(f) show the results of training a DT with the scattering coefficients of the hydrophone data for the looped and branched water networks models, respectively. The candidate set of leaf size values was leaf size = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50] for the looped and branched water network models. The Gini impurity was the split criterion to generate the DT. The vertical dashed black line shows the leaf size value that produces the best average validation accuracy, which was used to train the final DT and evaluate it with the test set.

On the other hand, Fig. 18 shows the ROC curves and the AUC of each tested machine-learning technique when evaluated with the test set for the looped and branched water network models. Table 6 summarizes the performance and results of each machine learning technique when evaluated with the test set in terms of accuracy, recall, precision, F1-score, and AUC. The hyperparameters that produce the

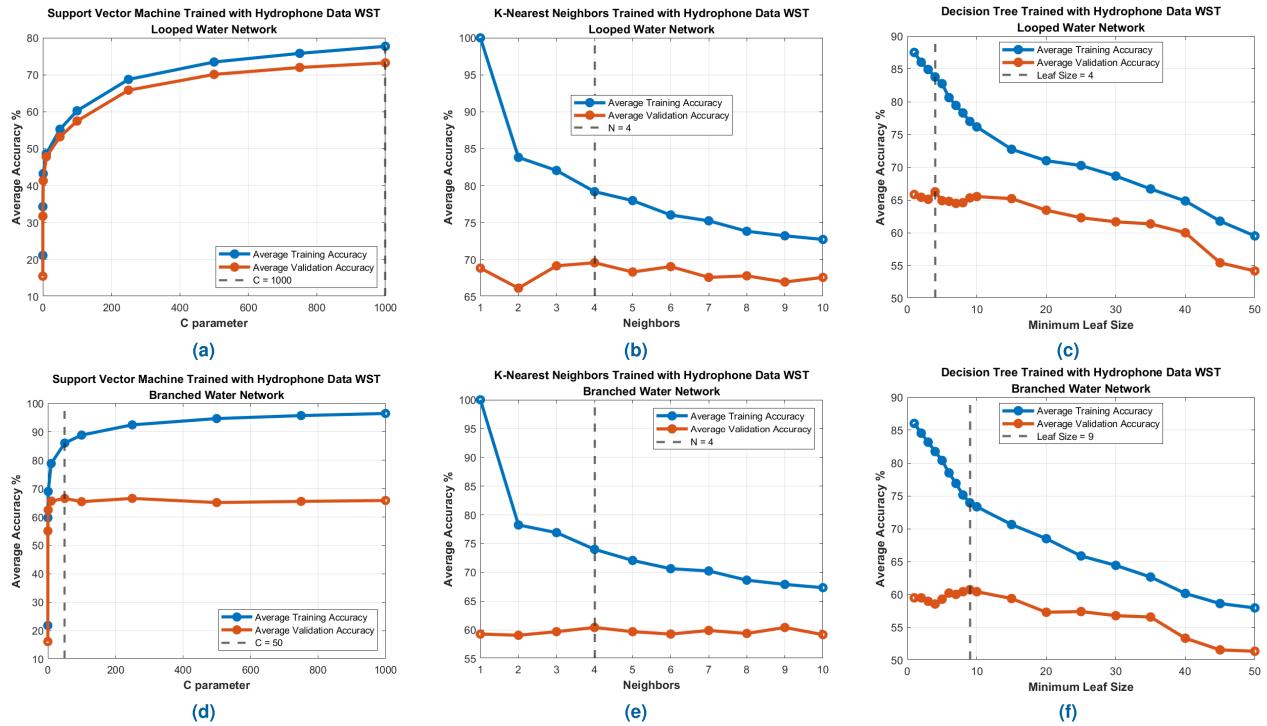


FIGURE 17. Training (blue) and validation (red) average accuracies while varying the hyperparameters of the SVM, KNN, and DT trained with the features of the hydrophone signals derived from the WST for water leakage classification. The vertical dashed line shows the hyperparameter value that produces the best trade-off between the training and validation sets' average accuracies. The top graphs show the results of the models trained for the looped water network; the bottom graphs show the results of models trained for the branched water network.

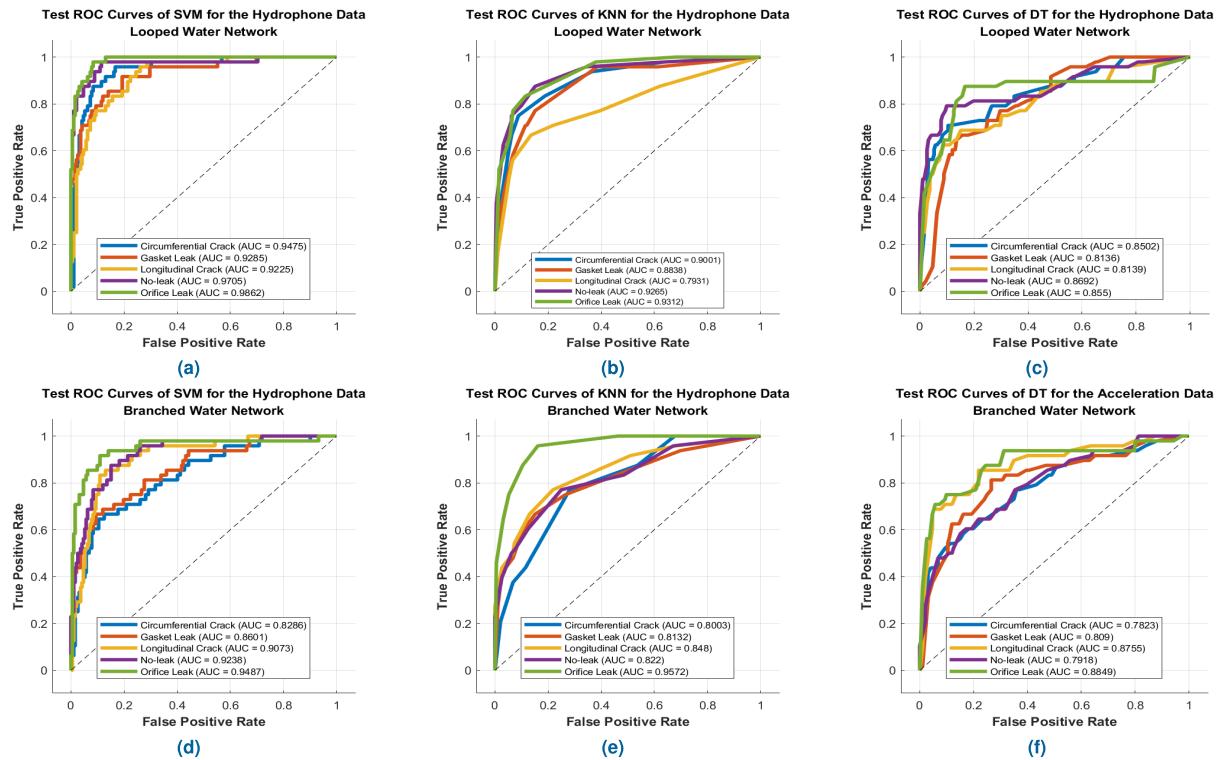


FIGURE 18. Test ROCs and AUCs of each machine learning technique trained for the looped (top) and branched (bottom) water networks.

reported performance are presented in Table 6; these hyperparameters coincide with the value of the hyperparameters

that produced the best average validation accuracy denoted by the vertical black dashed lines presented in Fig. 17.

TABLE 6. Results of testing the SVM, KNN, and DT trained with the scattering coefficients of the hydrophone data for water leakage classification for the looped and branched water network models.

Hydrophone Data										
Water Network	Machine Learning Technique	Parameters	Average Training Accuracy %	Average Validation Accuracy %	Testing Accuracy %	Water Leak Type %	Test Recall %	Test Precision %	Test F1-score %	Test AUC
Looped	SVM	Kernel: RBF Gamma: 1 C = 1000	77.7083 ± 0.9198	73.2292 ± 5.1043	78.3333	Circumferential Crack	79.1667	76	77.55	0.9475
						Gasket Leak	72.9167	72.9167	72.9167	0.9285
						Longitudinal Crack	68.75	73.3333	70.9677	0.9225
						No-leak	81.25	90.6977	85.7143	0.9705
						Orifice Leak	89.5833	79.6296	84.3137	0.9862
Branched	SVM	Kernel: RBF Gamma: 1 C = 50	86.0648 ± 0.6226	66.5625 ± 4.2084	70	Circumferential Crack	58.3333	63.6364	60.8696	0.8286
						Gasket Leak	60.4167	72.5	65.9091	0.8602
						Longitudinal Crack	75.000	64.2857	69.2308	0.9073
						No-leak	75.000	69.2308	72.000	0.9238
						Orifice Leak	81.25	81.25	81.25	0.9487
Looped	KNN	Neighbors: 4	79.2014 ± 0.6126	69.5833 ± 3.1366	68.75	Circumferential Crack	75	67.9245	71.2871	0.9001
						Gasket Leak	70.8333	60.7143	65.3846	0.8838
						Longitudinal Crack	56.25	67.5	61.3636	0.7931
						No-leak	75	75	75	0.9265
						Orifice Leak	66.6667	74.4186	70.3297	0.9312
Branched	KNN	Neighbors: 4	74.0046 ± 0.6949	60.4167 ± 4.8113	60	Circumferential Crack	60.4167	43.2836	50.4348	0.8003
						Gasket Leak	58.3333	59.5745	58.9474	0.8132
						Longitudinal Crack	54.1667	65	59.091	0.848
						No-leak	52.0833	62.5	56.8182	0.822
						Orifice Leak	75	78.2609	76.5957	0.9572
Looped	DT	Leaf Size: 4	83.7847 ± 0.4207	66.2500 ± 3.8415	64.1667	Circumferential Crack	60.4167	74.3590	66.6667	0.8502
						Gasket Leak	60.4167	55.7692	58.0000	0.8136
						Longitudinal Crack	62.5000	58.8235	60.6061	0.8139
						No-leak	72.9167	70.0000	71.4286	0.8692
						Orifice Leak	64.5833	64.5833	64.5833	0.855
Branched	DT	Leaf Size: 9	73.9468 ± 0.4954	60.7292 ± 4.6598	60.4167	Circumferential Crack	56.25	48.2143	51.9231	0.7823
						Gasket Leak	56.2500	56.2500	56.25	0.809
						Longitudinal Crack	68.75	76.7442	72.5275	0.8755
						No-leak	47.9167	57.5	52.2727	0.7918
						Orifice Leak	72.9167	66.0377	69.3069	0.8849

D. MULTIMODAL WATER LEAKAGE CLASSIFICATION

The multimodal model using LF was generated based on the results obtained for each data modality and each machine learning technique used to classify water leakages. According to the results presented in Tables 4, 5, and 6, the SVM was the technique that achieved the best test accuracy for each data modality and water network. Hence, the SVMs generated with the training set of the accelerometer, dynamic pressure, and hydrophone data WST were used to produce the ensemble of classifiers. Table 7 presents the best SVM models of each water network obtained for each data modality. Each trained classifier's predictions were combined with a majority voting mechanism when evaluated with the testing set. Based on this combination, the test accuracy, recall, precision, and F1-score were computed for each water leak type.

Figs. 19(a) and 19(b) present the confusion matrix of the LF models of the looped and branched water networks, respectively. Figs. 20(a) and 20(b) show the ROC curves of the LF models for each water leak type and the corresponding AUC of each ROC curve for the looped and branched

water networks, respectively. Finally, Table 8 summarizes the results of merging the predictions of the SVMs trained for each data modality and each water network model.

V. DISCUSSIONS

A. WATER LEAKAGE CLASSIFICATION PERFORMED VIA ACCELEROMETER DATA

Fig. 13 shows the training and validation curves obtained by applying 10-fold cross-validation to the SVM, KNN, and DT trained with the WST coefficients of the accelerometer data while varying the values of their respective hyperparameters for the looped and branched water network models. Based on these results, it is possible to appreciate that the SVM shows a lower degree of overfitting, and its performance increases with increasing values of C for both water network models. On the other hand, when using a low number of neighbors in the KNN algorithm, it tends to overfit, leading to a decrease in performance in terms of average validation accuracy. Moreover, the average validation accuracy further decreases as the number of neighbors increases. Similarly, the DT algorithm is prone to overfitting with low minimum leaf

TABLE 7. Trained models for each data modality and water network that were used to generate the LF model by combining their respective predictions using a majority voting mechanism.

Data Modality	Water Network	Machine Learning Technique	Parameters	Average Training Accuracy %	Average Validation Accuracy %	Testing Accuracy %
Vibration data	Looped	SVM	Kernel: RBF Gamma: 1 $C = 100$	90.4282 ± 0.3538	88.4375 ± 2.9646	89.5833
			Kernel: RBF Gamma: 1 $C = 750$	99.8958 ± 0.0366	97.7083 ± 1.5372	96.6667
	Branched	SVM	Kernel: RBF Gamma: 1 $C = 1000$	94.8495 ± 0.4092	88.8542 ± 3.4391	90
			Kernel: RBF Gamma: 1 $C = 1000$	87.6736 ± 0.8579	82.1875 ± 3.0449	84.5833
Dynamic Pressure data	Looped	SVM	Kernel: RBF Gamma: 1 $C = 1000$	77.7083 ± 0.9198	73.2292 ± 5.1043	78.3333
			Kernel: RBF Gamma: 1 $C = 50$	86.0648 ± 0.6226	66.5625 ± 4.2084	70
	Branched	SVM	Kernel: RBF Gamma: 1 $C = 1000$	95.8% 4.2%	95.8% 4.2%	77.1% 22.9%
			Kernel: RBF Gamma: 1 $C = 50$	100% 0.0%	97.9% 2.1%	100% 0.0%

size values. As the minimum leaf size values increase, the average validation accuracy of the DT decreases.

The results presented in Table 4 demonstrate that the SVM trained using the scattering coefficients of the accelerometer data achieves higher testing accuracy than the KNN and DT, for both water network topologies. In this regard, the SVM trained for the branched water network performs better than the one obtained for the looped network. It has been found that the proposed method is better at classifying water leakages in the branched water network than the looped water network when using accelerometer data as input. Similar results were observed for the KNN and DT algorithms, as their testing accuracy was higher for the branched water network model than for the looped water network model. Additionally, the SVM trained with the scattering coefficients of the accelerometer data performed better in identifying each water leak type. This is evident from the individual class F1-scores obtained compared to the DT and KNN algorithms. Nevertheless, for the looped water network SVM, the F1-score of the longitudinal crack water leakage was lower than that obtained with the KNN algorithm (see Table 4).

By analyzing the ROC curves of each of the classifiers trained with the scattering coefficients of the accelerometer data presented in Fig. 14, it is possible to observe that for each water leak type class, all of the classifiers are better than a random classifier since the ROC curves are closer to the top-left corner, which suggests an adequate classifier. Nevertheless, similar to the performance obtained in terms of testing accuracy and F1-score, the SVM classifiers for both water network types show a higher performance in terms of AUC for all of the water leak types compared to the KNN and DT. The lowest AUC obtained from the SVM trained for the looped water network was for the longitudinal crack water

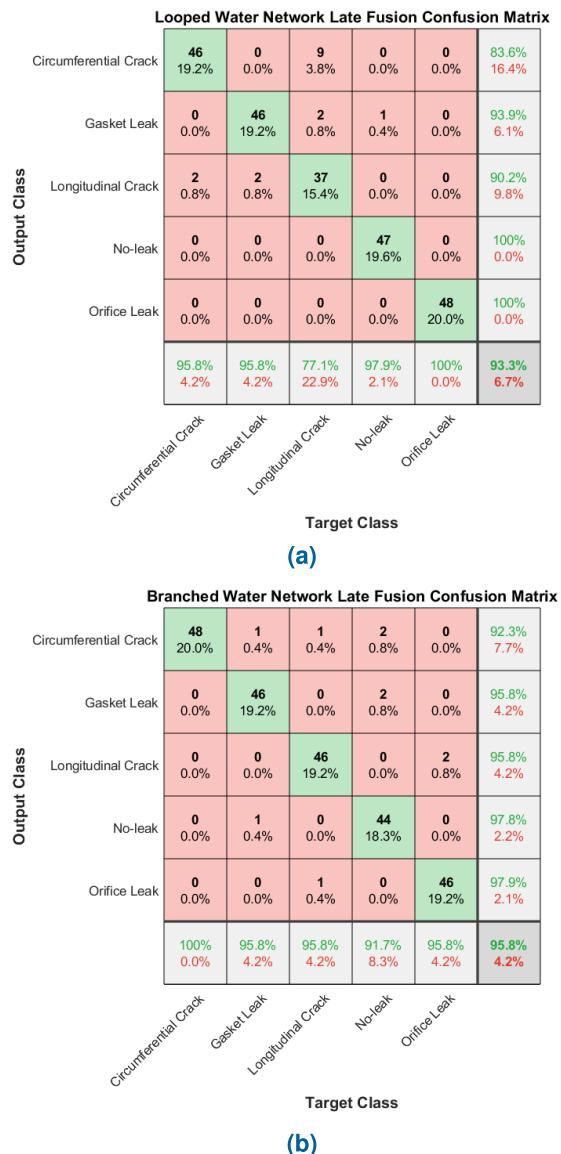


FIGURE 19. Confusion matrix of the LF models for the looped and branched water networks. a) Shows the confusion matrix of the looped water network. b) Shows the confusion matrix of the branched water network.

leakage, which is also the class that showed the lowest F1 score.

B. WATER LEAKAGE CLASSIFICATION PERFORMED VIA DYNAMIC PRESSURE DATA

Fig. 15 displays the training and validation curves obtained from the models trained with the dynamic pressure data scattering coefficients. The SVM, KNN, and DT were trained while varying their respective hyperparameters to generate the looped and branched water network models. In the case of the SVM, increasing the value of C leads to improved performance. Otherwise, with the KNN algorithm, lower neighbors' values lead to overfitting, and increasing the number of neighbors decreases its average validation

TABLE 8. Results of applying LF via a majority voting mechanism by considering the SMVs trained with the wavelet scattering coefficients of accelerometer, dynamic pressure, and hydrophone data for water leakage classification.

Water Network	Data Modality	Machine Learning Technique	Parameters	Test Accuracy %	Water Leak Type	Test Recall %	Test Precision %	Test F1-score %	Test AUC
Looped	Accelerometer	SVM	Kernel: RBF, Gamma: 1 C = 100	93.33	Circumferential Crack	95.8333	83.6364	89.3204	0.995
			Kernel: RBF, Gamma: 1 C = 1000		Gasket Leak	95.8333	93.8776	94.8454	0.9998
			Kernel: RBF, Gamma: 1 C = 1000		Longitudinal Crack	77.0833	90.2439	83.1461	0.9954
	Dynamic Pressure	SVM	Kernel: RBF, Gamma: 1 C = 1000		No-leak	97.9167	100	98.9474	0.999
			Kernel: RBF, Gamma: 1 C = 1000		Orifice Leak	100	100	100	1
			Kernel: RBF, Gamma: 1 C = 50		Circumferential Crack	100	92.3077	96	0.991
Branched	Accelerometer	SVM	Kernel: RBF, Gamma: 1 C = 750	95.833	Gasket Leak	95.8333	95.8333	95.8333	0.999
			Kernel: RBF, Gamma: 1 C = 1000		Longitudinal Crack	95.8333	95.8333	95.8333	0.9986
			Kernel: RBF, Gamma: 1 C = 50		No-leak	91.6667	97.7778	94.6237	0.9894
	Hydrophone	SVM	Kernel: RBF, Gamma: 1 C = 50		Orifice Leak	95.8333	97.8723	96.8421	0.9967
			Kernel: RBF, Gamma: 1 C = 50						
			Kernel: RBF, Gamma: 1 C = 50						

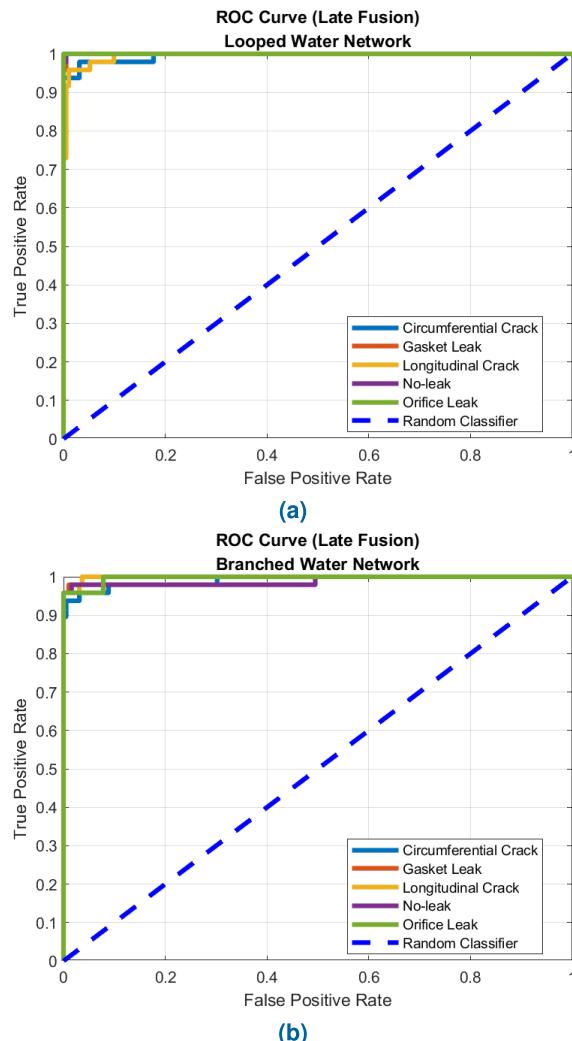


FIGURE 20. Test ROC curves and AUCs of the LF models. a) Shows the ROC curves of the LF model for the looped water network. b) Shows the ROC curves of the LF model for the branched water network.

accuracy. The best performance for the looped water network model is achieved with four neighbors, while the branched

water network is achieved with three neighbors. Similarly, the DT algorithm faces overfitting with lower minimum leaf sizes, and increasing the leaf size results in decreased average validation accuracy. The best performance of the DT for the looped water network model is achieved with a minimum leaf size of nine, while for the branched water network model, it's four.

Table 5 compares the average training and validation accuracies of the SVMs trained for the looped and branched water network models. In this regard, the testing accuracy of the SVMs is greater than that of the DT and KNN. The test F1-scores obtained for each water leak type from the SVM trained with the scattering coefficients of the dynamic pressure data were also greater than those obtained from the DT and KNN. However, for the circumferential crack F1-score, the KNN of the branched water network model yielded slightly better results than the SVM of the branched water network model.

The ROC curves obtained from evaluating the selected machine learning classifiers with the testing set, as shown in Fig. 16, demonstrate that they performed better than a random classifier. This is because all the ROC curves for all classes are closer to the top-left corner. However, the SVM classifier obtained a higher performance for each class and both water network models in terms of AUC.

C. WATER LEAKAGE CLASSIFICATION PERFORMED VIA HYDROPHONE DATA

Fig. 17 shows the training and validation curves of each machine learning technique trained with the scattering coefficients of the hydrophone data while varying their hyperparameters. In this regard, it is possible to observe that there is a certain degree of overfitting since there is a significant difference between the average training and validation accuracies for each tested machine learning technique, where for most of the hyperparameters, the average training accuracy is higher compared to the average validation accuracy. For the case of the SVM of the looped

water network, it is possible to observe that the performance increases when the regularization hyperparameter C increases. For the SVM trained for the branched water network, the model shows a greater difference between the average validation and training accuracy, with a maximum average validation accuracy achieved when $C = 50$. On the other hand, the KNN shows overfitting with low neighbor values, and its performance decreases when the number of neighbors used to make a prediction increases. Likewise, the DT shows overfitting with low leaf size values, and its average validation accuracy decreases when the minimum leaf size values increase.

The overfitting presented in each machine-learning technique influenced the lower testing accuracy of each technique, as shown in Table 6, where it can be appreciated that the obtained testing accuracy was below 80%. Despite the above, the performance of the SVM trained with the hydrophone data's scattering coefficients was higher than the other tested techniques. Regarding the F1-score, the SVMs achieved a higher performance in detecting each water leak type than the DT and KNN for both water network topologies.

By analyzing the ROC curves presented in Fig. 18, it can be observed that each curve is closer to the top left corner and is better than a random classifier, represented by the diagonal black dashed line. The SVM achieved an AUC greater than 0.9 for each looped water network model leak type, outperforming the AUC obtained by the DT and KNN. On the other hand, the branched water network has a greater variation in the AUC obtained from each trained model, where for some classes, the AUC is lower than 0.9.

D. COMPARISON BETWEEN DATA MODALITIES FOR WATER LEAKAGE CLASSIFICATION

After comparing the performance of different data modalities for water leakage classification, it was observed that the accelerometer data provided the highest testing accuracy for both types of water networks. The above is consistent with the relatively high performance that vibration data has shown in the literature (see Table 1) and its extensive use as discussed in [80]. Nevertheless, it is worth mentioning that the testing accuracy of the SVM of the looped water network trained with dynamic pressure data is slightly higher compared to the SVM of the looped water network trained with accelerometer data. By comparing the individual class performance in terms of the F1-score of the looped water network SVM for each data modality, it can be appreciated that the model trained with dynamic pressure data achieved a higher F1-score for the longitudinal and circumferential crack compared to the model trained with accelerometer data. Nonetheless, the SVM trained with the accelerometer data for the looped water network achieved a higher F1-score for most classes. In general, the SVMs trained with accelerometer data achieved a higher performance in terms of F1-score compared

to the ones trained with dynamic pressure and hydrophone data for both water networks.

On the other hand, the hydrophone data showed a lower performance when its scattering coefficients were used as input compared to the performance obtained from the dynamic pressure and accelerometer data. The above could be attributed to the background noise originating from the chainsaw and traffic sounds played simultaneously while the data was recorded (see Section III). Guo et al. [18] denoted the lack of effectiveness of using acoustic data to detect water leaks in the presence of noise. The above contrasts with the frequent use of acoustic data for water leak detection presented in Section II and Table 1.

In the case of pressure data, its performance was slightly better in terms of testing accuracy for the SVM of the looped water network compared to the SVM generated from accelerometer data. However, pressure data is commonly combined with flow data rather than individually, as presented in Section II. It can be observed from Tables 5, and 6 that dynamic pressure and hydrophone data perform better in classifying leaks on the looped water network than in the branched water network. On the contrary, the models trained with accelerometer data performed better in classifying leaks on the branched water network than in the looped water network. Table 7 summarizes the best models for each water network modality, their hyperparameters, and performance. These models were the ones used for the LF approach.

E. WATER LEAKAGE CLASSIFICATION VIA LATE FUSION

By performing the combination of the prediction of the SVMs trained for each data modality via LF (see Table 8 and Fig. 19), it is possible to observe that the testing accuracy for the looped water network model is higher compared to the testing accuracy of the models trained by considering only one type of data modality (see Table 7). By analyzing the individual class performance of performing LF, it can be observed that the performance in classifying orifice cracks achieved a perfect score in terms of recall, precision, F1-score, and AUC for the case of the looped water network model. The classification of no-leaks also showed an improvement in F1-score, precision, and AUC compared to only considering one data modality (see Tables 4, 5, and 6).

On the other hand, the branched water network LF model obtained a lower test accuracy than using only the accelerometer data as input (see Table 4). By comparing the individual class performance of the LF model of the branched water network, it can be observed that the performance obtained in terms of F1-score and AUC is better than that obtained by only using the hydrophone data and dynamic pressure data (see Tables 5 and 6). Furthermore, the ROC curves of the looped and branched water network models are shown in Figs. 20(a) and 20(b) demonstrate adequate classifiers since the ROC curve of each leak type is close to

the top-left corner, and in some classes, the curves overlap. Nonetheless, as previously commented, only the LF model of the looped water network achieved a better performance than the models trained with accelerometer data in terms of testing accuracy.

F. COMPARISON WITH THE LITERATURE

A comparison between the results presented in this study and those obtained in the related research is complicated. The above is mainly due to heterogeneity between the systems and water networks analyzed in each study; this aspect is also shown in Table 1. Thus, to make a homogenous comparison of the methods and results presented in the literature to the ones presented in this work, the machine learning algorithms should be trained and tested with the same dataset or data distribution. That is why different data modalities presented in this work were compared only with models trained and tested with Aghashahi et al. [53] dataset. Furthermore, to the authors' knowledge, no other studies are available in the literature in which Aghashahi's dataset was used for water leakage classification based on machine learning, with which a comparison can be made.

Despite this, a brief comparison is presented regarding the accuracy of the related research. By analyzing Table 1, it is possible to observe that accuracy is the most common metric employed in the literature. In this regard, the lowest accuracy related to water leakage detection is the study of Fares et al. [24] with an accuracy of 87.2%. On the other hand, the study of Ullah et al. [26] reported an accuracy of 100%. In addition, these two studies employed neural networks as classifiers, which are frequently used in the context of water leakage detection [12], [18], [19], [20], [22], [31], [37], [81]. In this regard, the results reported by the LF model of the looped and branched water network models are within those minimum and maximum values with a testing accuracy of 93.33% and 95.833%, respectively (see Table 8). On the other hand, the testing accuracy of only employing the scattering coefficients of the accelerometer data is also within the minimum and maximum accuracy values reported in the literature (see Table 4). In the case of the hydrophone data, the models of both water networks attained lower accuracies (see Table 6). However, the classes considered in this study and the related research also differ since most of the literature has been concerned with tackling the problem as a binary classification problem, detecting the presence or absence of the leak [18], [19], [23], [25], [43], [44], [46]. In contrast, this study aimed to determine the type of leak on a water pipe, which makes the classification problem more complex as it involves multiple classes rather than just two. Therefore, it is more appropriate to compare the performance of each study by evaluating the sensitivity achieved by each class, rather than the accuracy, which can be biased as it is an overall metric of TP and TN. Nonetheless, comparing studies is challenging since authors have not reported the same metrics, and the classes differ between studies.

The main difference between this study and the related research is using the WST as a feature extraction technique. Despite that time, frequency, and time-frequency representations have been widely used for feature extraction, the features that have been computed vary among the studies (see Section II). In contrast, the WST seeks to create a translation invariant representation of the signal of interest that is stable to small deformations to reduce the within-class variance [45], [59]. In this sense, a feature extraction process is applied to the data without guessing features or relying on deep learning [81]. Furthermore, LF was used to simultaneously consider the predicting power of accelerometer, pressure, and hydrophone data for water leakage classification. This aspect remains underexplored in the literature and has been limited to only comparing and considering pressure and flow data [7], [13], [27], [29], [30].

VI. LIMITATIONS

Despite the relatively high performance obtained by using the scattering transform of accelerometer data for classifying water leakages and the performance attained via the LF models, certain limitations should be considered. The first one is related to the fact that the response of the sensors to water leakages was recorded under laboratory conditions and consisted of a scale water network that cannot be compared to the complex water networks that cities and certain households could possess. Moreover, the related research has been concerned with testing data-driven techniques for water leak detection in cities' water networks such as in [18], [20], [21], [22], and [46], an aspect that was not addressed in the present research but can be explored in future works to analyze the performance of the proposed method when tested on more complex water networks. Furthermore, the leak position was set in a predefined position, with no variation in its location. Hence, leak localization was not addressed in this study, but has been considered in related research as in [82].

A potential bias source is the location of the sensors and the number of sensors considered for each data modality. In this work, the sensor configuration and placement shown in Fig. 2 produced the results presented in Section IV. However, it cannot be accepted that the proposed sensor placement is the most adequate for classifying water leakages for both water network topologies. In addition, a different sensor placement could produce a different response on the sensors used to analyze water leakages, which could impact the performance of the proposed method. Additionally, the pipe materials (e.g., metallic or plastic pipes) are another factor that could alter the sensor's response, which could affect the performance of the feature extraction method and training of the classifiers [44], [47]. Thus, the data distribution may affect the training and performance of data-driven techniques; a limitation also commented on by McMillan et al. [61] when performing water leakage detection.

Another point that was not addressed was the comparison of the WST features with the traditional feature extraction

methods used in the literature to show the effectiveness of the proposed method. However, the above requires an in-depth analysis of the appropriate features to perform this task since, as stated in Section II-F, the related research has focused on leak detection only (i.e., binary classification) and not on the classification of different leak types. Hence, an appropriate feature extraction should be performed for each leak type considered in this study since it cannot be ensured that the features used for leak detection could be adequate to classify different leak types. The above could allow a fair comparison between the WST and traditional feature extraction methods for water leakage classification.

Another limitation of this study is how the parameters of the WST were selected. In this study, the selection of the number of voices per octave of each scattering layer was set to $Q = 1$, and the invariance scale was set to 0.01 seconds, which were selected based on the number of input features that it produced, which were 20 for the case of the accelerometer and dynamic pressure data, and 9 for the hydrophone data since other parametrizations produce a higher number of scattering coefficients, which could produce overfitting on the trained models. Hence, a methodology that allows for a better selection of the WST parameterization that could improve the performance in classifying water leakages could be proposed. Another omitted aspect was the application of feature selection on the obtained scattering coefficients, which could impact the performance of the selected machine learning techniques.

Related to how the machine learning techniques' hyperparameters were selected, this study set one of their hyperparameters and varied only one to determine the best performance hyperparameters. For example, in the case of the SVM, the gamma hyperparameter was fixed to 1, and the C hyperparameter was varied to determine an adequate value. However, this hyperparameter selection does not assure that its selection is the best for training the model. Thus, other methods, such as Grid Search, could provide a better selection of the selected machine learning techniques' hyperparameters that improve the overall methodology's performance [83].

In the case of the LF approach, one of its main disadvantages is that a model for each data modality should be considered, increasing the overall model's complexity. On the contrary, other multimodal approaches, such as early fusion, only required a single model. Thus, another aspect that remains to be explored is the comparison of different multimodal techniques for performing water leakage classification [84]. Another bias source of the LF approach presented in this study is the assumption that the sampling of the acoustic, accelerometer, and pressure data was performed synchronously. However, the dataset documentation does not explicitly confirm this assumption. Hence, it is difficult to mitigate this issue with the available information presented in [53] and [54].

Additionally, machine learning classifiers were combined with the scattering coefficients of the accelerometer,

hydrophone, and dynamic pressure data for water leakage classification. This was validated using a cross-validation approach on a publicly available dataset. Nonetheless, implementing the proposed method in a real set-up was not carried out, which could bring out challenges such as adequate hardware selection and the need to evaluate the method under real conditions. The above influences the potential real-time or online execution of the proposed method since, to select an appropriate hardware device to execute the proposed method, aspects such as the computation time needed to classify the water leakage properly should be established effectively. The selected hardware device needed to implement the proposed method should be capable of recording a window of 1 second for each sensor type simultaneously, filtering the data through wavelet denoising, executing the WST, feeding the generated features to the selected machine learning classifiers to classify the data in the different leak types, and executing the classifier. The hardware implementation of a water leakage detection algorithm is an aspect that has been addressed in related research, such as in the work of Islam et al. [85], but this work did not consider it. Hence, there is also the opportunity to evaluate the proposed method in a physical set-up rather than limiting the analysis to simulation results.

Finally, in this study, the classification of four leak types aside from no-leaks was performed. On the other hand, related studies have proposed to classify the size of the leak, such as in [26], [43], and [64]. This aspect was not considered in the present study due to how the data was collected in the work of Aghashahi et al. [53], but the leak size may degrade the performance of the WST and the trained classifiers. Accordingly, future work could explore the effect of the leak size when employing the WST as a feature extraction technique of hydrophone, accelerometer, and pressure data for water leakage classification.

VII. CONCLUSION AND FUTURE WORK

This work proposed using the WST as a feature extraction technique of acceleration, dynamic pressure, and acoustic data for water leakage classification by employing an SVM, KNN, and DT. The performance of each data modality for water leakage classification was compared, and later, the prediction of the best models of each data modality was combined using an LF approach. The results of this study show that the use of the scattering coefficients of accelerometer data and an SVM obtained the best performance in classifying water leakages in the two water network topologies considered in this study with a testing accuracy of 89.5833% for the looped water network model and 96.6667% for the branched water network model. The LF model results for the looped water network achieved a higher accuracy than the unimodal classifiers, with a test accuracy of 93.33%. Finally, the LF model of the branched network had a lower test accuracy than the SVM trained only with the accelerometer data WST.

Future studies could explore improving the performance of water leakage classification using machine learning techniques and the WST. For example, one aspect not addressed in this work is the variation in the parametrization of the WST, which should be tested to evaluate its impact on classification accuracy. Another possible improvement is using feature selection to identify the most relevant scattering coefficients computed from the accelerometer, dynamic pressure, and hydrophone data. Furthermore, multimodal techniques such as early and hybrid fusion could be tested to determine the best approach to integrating different data types for water leakage classification. Despite the relatively high accuracy that the WST in combination with SVMs attained, the models were trained and tested from data collected from a laboratory testbed. Hence, it is crucial to evaluate the performance of the methods used in this study on a real set-up or water distribution system. Finally, a study of the features derived from the WST compared to traditional feature extraction methods can be performed to determine, under equal conditions, the effectiveness of the WST as a feature extraction technique for water leakage classification based on acceleration, acoustic, and dynamic pressure data.

REFERENCES

- [1] Y. Liu, X. Ma, Y. Li, Y. Tie, Y. Zhang, and J. Gao, "Water pipeline leakage detection based on machine learning and wireless sensor networks," *Sensors*, vol. 19, no. 23, p. 5086, Nov. 2019.
- [2] K. Xin, F. Li, T. Tao, N. Xiang, and Z. Yin, "Water losses investigation and evaluation in water distribution system—The case of SA city in China," *Urban Water J.*, vol. 12, no. 5, pp. 430–439, Jun. 2014.
- [3] F. Wang, W. Lin, Z. Liu, S. Wu, and X. Qiu, "Pipeline leak detection by using time-domain statistical features," *IEEE Sensors J.*, vol. 17, no. 19, pp. 6431–6442, Oct. 2017.
- [4] A. Lay-Ekuakille and P. Vergallo, "Decimated signal diagonalization method for improved spectral leak detection in pipelines," *IEEE Sensors J.*, vol. 14, no. 6, pp. 1741–1748, Jun. 2014.
- [5] X. Fan and X. Yu, "An innovative machine learning based framework for water distribution network leakage detection and localization," *Struct. Health Monitor.*, vol. 21, no. 4, pp. 1626–1644, Aug. 2021.
- [6] A. Cataldo, G. Cannazza, E. De Benedetto, and N. Giaquinto, "A new method for detecting leaks in underground water pipelines," *IEEE Sensors J.*, vol. 12, no. 6, pp. 1660–1667, Jun. 2012.
- [7] H. M. Tornyeviadzi and R. Seidu, "Leakage detection in water distribution networks via 1D CNN deep autoencoder for multivariate SCADA data," *Eng. Appl. Artif. Intell.*, vol. 122, Jun. 2023, Art. no. 106062.
- [8] T. Platte, "Evaluation of the uncertainty of measurement for a dynamic calibration approach for pressure sensors," *Meas., Sensors*, vol. 18, Dec. 2021, Art. no. 100252.
- [9] A. Moubayed, M. Sharif, M. Luccini, S. Primak, and A. Shami, "Water leak detection survey: Challenges & research opportunities using data fusion & federated learning," *IEEE Access*, vol. 9, pp. 40595–40611, 2021.
- [10] D. Zaman, M. K. Tiwari, A. K. Gupta, and D. Sen, "A review of leakage detection strategies for pressurised pipeline in steady-state," *Eng. Failure Anal.*, vol. 109, Jan. 2020, Art. no. 104264.
- [11] D. Barrientos-Torres, E. A. Martinez-Ríos, S. A. Navarro-Tuch, J. L. Pablos-Hach, and R. Bustamante-Bello, "Water flow modeling and forecast in a water branch of Mexico city through ARIMA and transfer function models for anomaly detection," *Water*, vol. 15, no. 15, p. 2792, Aug. 2023.
- [12] J. Bohorquez, M. F. Lambert, B. Alexander, A. R. Simpson, and D. Abbott, "Stochastic resonance enhancement for leak detection in pipelines using fluid transients and convolutional neural networks," *J. Water Resour. Planning Manage.*, vol. 148, no. 3, Mar. 2022, Art. no. 04022001.
- [13] Z. Hu, B. Chen, W. Chen, D. Tan, and D. Shen, "Review of model-based and data-driven approaches for leak detection and location in water distribution systems," *Water Supply*, vol. 21, no. 7, pp. 3282–3306, Apr. 2021.
- [14] S. R. Mounce, A. J. Day, A. S. Wood, A. Khan, P. D. Widdop, and J. Machell, "A neural network approach to burst detection," *Water Sci. Technol.*, vol. 45, nos. 4–5, pp. 237–246, Feb. 2002.
- [15] M. R. Islam, S. Azam, B. Shanmugam, and D. Mathur, "A review on current technologies and future direction of water leakage detection in water distribution network," *IEEE Access*, vol. 10, pp. 107177–107201, 2022.
- [16] P. Mohan Doss, M. M. Rokstad, D. Steffelbauer, and F. Tscheikner-Gratl, "Uncertainties in different leak localization methods for water distribution networks: A review," *Urban Water J.*, vol. 20, no. 8, pp. 953–967, Jul. 2023.
- [17] S. G. Vrachimis, D. G. Eliades, R. Taormina, Z. Kapelan, A. Ostfeld, S. Liu, M. Kyriakou, P. Pavlou, M. Qiu, and M. M. Polycarpou, "Battle of the leakage detection and isolation methods," *J. Water Resour. Planning Manage.*, vol. 148, no. 12, Dec. 2022, Art. no. 04022068.
- [18] G. Guo, X. Yu, S. Liu, Z. Ma, Y. Wu, X. Xu, X. Wang, K. Smith, and X. Wu, "Leakage detection in water distribution systems based on time-frequency convolutional neural network," *J. Water Resour. Planning Manage.*, vol. 147, no. 2, Feb. 2021, Art. no. 04020101.
- [19] Z. Fereidooni, H. Tahayori, and A. Bahadori-Jahromi, "A hybrid model-based method for leak detection in large scale water distribution networks," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 2, pp. 1613–1629, Jul. 2020.
- [20] R. Vanijjirattikhun, S. Khomsay, N. Kitbutrawat, K. Khomsay, U. Supakchukul, S. Udomsuk, J. Suwatthikul, N. Oumtrakul, and K. Anusart, "AI-based acoustic leak detection in water distribution systems," *Results Eng.*, vol. 15, Sep. 2022, Art. no. 100557.
- [21] Y. Shen and W. Cheng, "A tree-based machine learning method for pipeline leakage detection," *Water*, vol. 14, no. 18, p. 2833, Sep. 2022.
- [22] T. Yu, X. Chen, W. Yan, Z. Xu, and M. Ye, "Leak detection in water distribution systems by classifying vibration signals," *Mech. Syst. Signal Process.*, vol. 185, Feb. 2023, Art. no. 109810.
- [23] D. P. Sousa, R. Du, J. M. B. da Silva, C. C. Cavalcante, and C. Fischione, "Leakage detection in water distribution networks using machine-learning strategies," *Water Supply*, vol. 23, no. 3, pp. 1115–1126, Feb. 2023.
- [24] A. Fares, I. A. Tijani, Z. Rui, and T. Zayed, "Leak detection in real water distribution networks based on acoustic emission and machine learning," *Environ. Technol.*, vol. 44, no. 25, pp. 3850–3866, May 2022.
- [25] M. Saravanabalaji, N. Sivakumaran, S. Ranganthan, and V. Athappan, "Acoustic signal based water leakage detection system using hybrid machine learning model," *Urban Water J.*, vol. 20, no. 9, pp. 1123–1139, Aug. 2023.
- [26] N. Ullah, Z. Ahmed, and J.-M. Kim, "Pipeline leakage detection using acoustic emission and machine learning algorithms," *Sensors*, vol. 23, no. 6, p. 3226, Mar. 2023.
- [27] M. Kammoun, A. Kammoun, and M. Abid, "Experiments based comparative evaluations of machine learning techniques for leak detection in water distribution systems," *Water Supply*, vol. 22, no. 1, pp. 628–642, Aug. 2021.
- [28] E. Sahin and H. Yüce, "Prediction of water leakage in pipeline networks using graph convolutional network method," *Appl. Sci.*, vol. 13, no. 13, p. 7427, Jun. 2023.
- [29] N. Mashhadi, I. Shahrouz, N. Attoue, J. El Khattabi, and A. Aljer, "Use of machine learning for leak detection and localization in water distribution systems," *Smart Cities*, vol. 4, no. 4, pp. 1293–1315, Oct. 2021.
- [30] V. Tyagi, P. Pandey, S. Jain, and P. Ramachandran, "A two-stage model for data-driven leakage detection and localization in water distribution networks," *Water*, vol. 15, no. 15, p. 2710, Jul. 2023.
- [31] X. Hu, Y. Han, B. Yu, Z. Geng, and J. Fan, "Novel leakage detection and water loss management of urban water supply network using multiscale neural networks," *J. Cleaner Prod.*, vol. 278, Jan. 2021, Art. no. 123611.
- [32] S. M. Bazaz, M. Lohtander, and J. Varis, "The prediction method of tool life on small lot turning process—Development of digital twin for production," *Proc. Manuf.*, vol. 51, pp. 288–295, Oct. 2020.
- [33] J. Choi and S. Im, "Application of CNN models to detect and classify leakages in water pipelines using magnitude spectra of vibration sound," *Appl. Sci.*, vol. 13, no. 5, p. 2845, Feb. 2023.
- [34] Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: A big data–AI integration perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 4, pp. 1328–1347, Apr. 2021.

- [35] D. Karimanzira, "Simultaneous pipe leak detection and localization using attention-based deep learning autoencoder," *Electronics*, vol. 12, no. 22, p. 4665, Nov. 2023.
- [36] C.-C. Jin and X. Chen, "An end-to-end framework combining time-frequency expert knowledge and modified transformer networks for vibration signal classification," *Expert Syst. Appl.*, vol. 171, Jun. 2021, Art. no. 114570.
- [37] Y. Wu, X. Ma, G. Guo, Y. Huang, M. Liu, S. Liu, J. Zhang, and J. Fan, "Hybrid method for enhancing acoustic leak detection in water distribution systems: Integration of handcrafted features and deep learning approaches," *Process Saf. Environ. Protection*, vol. 177, pp. 1366–1376, Sep. 2023.
- [38] J. D. Butterfield, A. Krynnin, R. P. Collins, and S. B. M. Beck, "Experimental investigation into vibro-acoustic emission signal processing techniques to quantify leak flow rate in plastic water distribution pipes," *Appl. Acoust.*, vol. 119, pp. 146–155, Apr. 2017.
- [39] J. Sun, Q. Xiao, J. Wen, and Y. Zhang, "Natural gas pipeline leak aperture identification and location based on local mean decomposition analysis," *Measurement*, vol. 79, pp. 147–157, Feb. 2016.
- [40] K. T. Aminu, D. McGlinchey, and A. Cowell, "Acoustic signal processing with robust machine learning algorithm for improved monitoring of particulate solid materials in a gas flowline," *Flow Meas. Instrum.*, vol. 65, pp. 33–44, Mar. 2019.
- [41] Z. Xue, L. Tao, J. Fuchun, E. Riehle, H. Xiang, N. Bowen, and R. P. Singh, "Application of acoustic intelligent leak detection in an urban water supply pipe network," *J. Water Supply, Res. Technol.-Aqua*, vol. 69, no. 5, pp. 512–520, Jun. 2020.
- [42] M. Ahadi and M. S. Bakhtiar, "Leak detection in water-filled plastic pipes through the application of tuned wavelet transforms to acoustic emission signals," *Appl. Acoust.*, vol. 71, no. 7, pp. 634–639, Jul. 2010.
- [43] R. Zese, E. Bellodi, C. Luciani, and S. Alvisi, "Neural network techniques for detecting intra-domestic water leaks of different magnitude," *IEEE Access*, vol. 9, pp. 126135–126147, 2021.
- [44] S. Lee and B. Kim, "Machine learning model for leak detection using water pipeline vibration sensor," *Sensors*, vol. 23, no. 21, p. 8935, Nov. 2023.
- [45] J. Andén and S. Mallat, "Deep scattering spectrum," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4114–4128, Aug. 2014.
- [46] T. Ravichandran, K. Gavahi, K. Ponnambalam, V. Burtea, and S. J. Mousavi, "Ensemble-based machine learning approach for improved leak detection in water mains," *J. Hydroinformatics*, vol. 23, no. 2, pp. 307–323, Jan. 2021.
- [47] S. Tariq, B. Bakhtawar, and T. Zayed, "Data-driven application of MEMS-based accelerometers for leak detection in water distribution networks," *Sci. Total Environ.*, vol. 809, Feb. 2022, Art. no. 151110.
- [48] H. Kim, H. Nam, W. Jung, and J. Lee, "Performance analysis of CNN frameworks for GPUs," in *Proc. IEEE Int. Symp. Perform. Anal. Syst. Softw. (ISPASS)*, Apr. 2017, pp. 55–64.
- [49] S. Morales and M. E. Bowers, "Time-frequency analysis methods and their application in developmental EEG data," *Develop. Cognit. Neurosci.*, vol. 54, Apr. 2022, Art. no. 101067.
- [50] O. Atanane, A. Mourhir, N. Benamar, and M. Zennaro, "Smart buildings: Water leakage detection using TinyML," *Sensors*, vol. 23, no. 22, p. 9210, Nov. 2023.
- [51] L. Mei, J. Zhou, S. Li, M. Cai, and T. Li, "Leak identification based on CS-ResNet under different leakage apertures for water-supply pipeline," *IEEE Access*, vol. 10, pp. 57783–57795, 2022.
- [52] E. A. Martinez-Ríos, R. Bustamante-Bello, and S. A. Navarro-Tuch, "Generalized Morse wavelets parameter selection and transfer learning for pavement transverse cracking detection," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106355.
- [53] M. Aghashahi, L. Sela, and M. K. Banks, "Benchmarking dataset for leak detection and localization in water distribution systems," *Data Brief*, vol. 48, Jun. 2023, Art. no. 109148.
- [54] M. Aghashahi, L. Sela, and M. K. Banks, "Dataset of leak simulations in experimental testbed water distribution system," *Mendeley Data*, VI, 2022, doi: [10.17632/tbrnp6vrnj.1](https://doi.org/10.17632/tbrnp6vrnj.1).
- [55] S. Mallat, "Group invariant scattering," *Commun. Pure Appl. Math.*, vol. 65, no. 10, pp. 1331–1398, Jul. 2012.
- [56] E. A. Martinez-Ríos, R. Bustamante-Bello, S. Navarro-Tuch, and H. Perez-Meana, "Applications of the generalized Morse wavelets: A review," *IEEE Access*, vol. 11, pp. 667–688, 2023.
- [57] E. Martinez-Ríos, L. Montesinos, and M. Alfaro-Ponce, "A machine learning approach for hypertension detection based on photoplethysmography and clinical data," *Comput. Biol. Med.*, vol. 145, Jun. 2022, Art. no. 105479.
- [58] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 2009.
- [59] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1872–1886, Aug. 2013.
- [60] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. Cham, Switzerland: Springer, 2021.
- [61] L. McMillan, J. Fayaz, and L. Varga, "Domain-informed variational neural networks and support vector machines based leakage detection framework to augment self-healing in water distribution networks," *Water Res.*, vol. 249, Feb. 2024, Art. no. 120983.
- [62] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A brief review of nearest neighbor algorithm for learning and classification," in *Proc. Int. Conf. Intell. Comput. Control Syst. (ICCS)*, May 2019, pp. 1255–1260.
- [63] M. Du, N. Liu, and X. Hu, "Techniques for interpretable machine learning," *Commun. ACM*, vol. 63, no. 1, pp. 68–77, Dec. 2019.
- [64] M. A. Virk, M. F. Mysorewala, L. Cheded, and I. M. Ali, "Leak detection using flow-induced vibrations in pressurized wall-mounted water pipelines," *IEEE Access*, vol. 8, pp. 188673–188687, 2020.
- [65] W. S. Noble, "What is a support vector machine?" *Nature Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006.
- [66] A. Subasi, K. Khateeb, T. Brahimi, and A. Sarirete, "Human activity recognition using machine learning methods in a smart healthcare environment," in *Innovation in Health Informatics*. New York, NY, USA: Academic, 2020, pp. 123–144.
- [67] S. B. Kotsiantis, "Decision trees: A recent overview," *Artif. Intell. Rev.*, vol. 39, no. 4, pp. 261–283, Jun. 2011.
- [68] Q. Pan, L. Zhang, G. Dai, and H. Zhang, "Two denoising methods by wavelet transform," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3401–3406, Nov. 1999.
- [69] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [70] G. Chen, W. Xie, and Y. Zhao, "Wavelet-based denoising: A brief review," in *Proc. 4th Int. Conf. Intell. Control Inf. Process.*, Jun. 2013, pp. 570–574.
- [71] M. Andreux, T. Angles, G. Exarchakis, R. Leonardiuzzi, G. Rochette, L. Thiry, J. Zarka, S. Mallat, J. Andan, E. Belilovsky, J. Bruna, V. Lostanlen, M. Chaudhary, M. J. Hirn, E. Oyallon, S. Zhang, C. Cellia, and M. Eickenberg, "Kymatio: Scattering transforms in Python," *J. Mach. Learn. Res.*, vol. 21, no. 60, pp. 1–6, 2020.
- [72] T. Baltrusaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 423–443, Feb. 2019.
- [73] S. C. Huang et al., "Fusion of medical imaging and electronic health records using deep learning: A systematic review and implementation guidelines," *Npj Digit. Med.*, vol. 3, p. 136, 2020, doi: [10.1038/s41746-020-00341-z](https://doi.org/10.1038/s41746-020-00341-z).
- [74] A. Barua, M. U. Ahmed, and S. Begum, "A systematic literature review on multimodal machine learning: Applications, challenges, gaps and future directions," *IEEE Access*, vol. 11, pp. 14804–14831, 2023.
- [75] S. M. Malakouti, "Babysitting hyperparameter optimization and 10-fold-cross-validation to enhance the performance of ML methods in predicting wind speed and energy generation," *Intell. Syst. Appl.*, vol. 19, Sep. 2023, Art. no. 200248.
- [76] F. Melo, *Area Under the ROC Curve*. Cham, Switzerland: Springer, 2013, pp. 38–39.
- [77] F. Pedregosa, S. Varoquaux, A. Gramfort, V. Michel, and B. Thirion, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Dec. 2011.
- [78] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Cham, Switzerland: Springer, 2009.
- [79] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Evanston, IL, USA: Routledge, Oct. 2017.
- [80] M. I. Mohd Ismail, R. A. Dzilyauddin, N. A. Ahmad Salleh, F. Muhammad-Sukki, N. Aini Bani, M. A. Mohd Izhar, and L. A. Latiff, "A review of vibration detection methods using accelerometer sensors for water pipeline leakage," *IEEE Access*, vol. 7, pp. 51965–51981, 2019.

- [81] U. Rajasekaran and M. Kothandaraman, "A novel custom one-dimensional time-series DenseNet for water pipeline leak detection and localization using acousto-optic sensor," *IEEE Access*, vol. 12, pp. 7966–7973, 2024.
- [82] I. Lucin, Z. Carija, S. Družeta, and B. Lucin, "Detailed leak localization in water distribution networks using random forest classifier and pipe segmentation," *IEEE Access*, vol. 9, pp. 155113–155122, 2021.
- [83] H. Alibrahim and S. A. Ludwig, "Hyperparameter optimization: Comparing genetic algorithm against grid search and Bayesian optimization," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jun. 2021, pp. 1551–1559.
- [84] P. Rashinkar and V. S. Krushnasamy, "An overview of data fusion techniques," in *Proc. Int. Conf. Innov. Mech. Ind. Appl. (ICIMIA)*, Feb. 2017, pp. 694–697.
- [85] M. R. Islam, S. Azam, B. Shanmugam, and D. Mathur, "An intelligent IoT and ML-based water leakage detection system," *IEEE Access*, vol. 11, pp. 123625–123649, 2023.



DAVID BARRIENTOS was born in Mexico City, Mexico, in August 1995. He is currently pursuing the Ph.D. degree in engineering sciences with Tecnológico de Monterrey, conducting research and disseminating knowledge about smart water management in the context of smart cities. His research interests include water scarcity mitigation through leakage assessment and efficient management, hydroinformatics, and data science.



ROGELIO BUSTAMANTE was born in Orizaba, Veracruz, in 1965. He received the B.S. degree in electronic instrumentation from the Faculty of Physics and Mathematics, Universidad Veracruzana, Xalapa, Veracruz, the master's degree in electronics and telecommunications from CICESE, Ensenada, Baja California, México, and the Ph.D. degree in telecommunications and electronics from SEPI, Instituto Politécnico Nacional. He has worked for more than 12 years in advanced bioinstrumentation systems, applying them to the driving assistance systems (ADAS) and the development of healthy active spaces (emotional domotics) and exoskeletons regarding the movement of lower limbs and bioinstrumentation systems. He has filed multiple patents with IMPI and WIPO.



ERICK AXEL MARTINEZ-RÍOS was born in Mexico City, Mexico, in May 1995. He received the B.S. degree (Hons.) in mechatronics engineering and the M.S. degree (Hons.) in engineering science from Instituto Tecnológico y de Estudios Superiores de Monterrey named as (Tecnológico de Monterrey), Mexico City, in 2017 and 2022, respectively, where he is currently pursuing the Ph.D. degree in engineering science. From 2018 to 2020, he was a Project Specialist with Instituto Tecnológico y de Estudios Superiores de Monterrey. During this stay, he worked on research and development projects in the field of robotics. He is the author of eight journal articles, two conference papers, and one book chapter. His research interests include signal processing, machine learning, and control theory.