Comprehensive Framework for Leak Detection in Water Distribution Systems: A Vibration Signal Processing and Machine Learning Approach

Hayder Wthaij Ajeel Al Ghasheem [1,2]
[1]Islamic Azad University, Science and Research Branch, Iran
[2]Ministry of water Resources, Baghdad, Iraq
Email: haider2eng222@gmail.com

Article History

Received: Mar. 12, 2025

Revised: May 22, 2025

Accepted: May 28, 2025

Abstract

This study provides a new approach for identifying water distribution system leaks. It blends vibration signal processing with machine learning. The system is based on vibration signals from accelerometers for pipeline observation through non-invasive methods and real-time. Based on a Random Forest classifier, the system is able to differentiate between different leak scenarios from no-leak cases with an accuracy of 97.3%. We validated the findings using a confusion matrix, which confirmed some cases of misclassification, indicating there is still much scope for improvement. We identified key statistical features such as RMS, kurtosis, and variance as being of prime importance for leak identification using feature importance analysis. These features enable capturing the specific vibration patterns of diverse leaks, allowing for accurate identification. This is an improvement over conventional leak detection techniques, offering a more reliable and efficient method for pipeline observation. The study also discusses how the procedure could make water distribution systems sustainable and operationally efficient for application in the real world.

Keywords- Pipeline leak detection, Vibration signal processing, Machine learning, Random Forest.

I. INTRODUCTION

Pipeline networks are essential for transporting water, oil, gas, and other fluids over long distances, and any undetected leak can lead to serious economic, environmental, and safety consequences [1]. Traditional leak-detection techniques—such as pressure monitoring, acoustic sensing, and manual inspection—often struggle under changing flow conditions and can miss small or intermittent leaks [2,3]. Vibration-based monitoring, by contrast, offers a non-invasive, real-time window into pipeline integrity: accelerometers mounted on the pipe detect subtle oscillations caused by fluid escaping through a defect [4] Recent work has shown that combining statistical analysis of vibration signals with machine learning significantly boosts detection accuracy. For instance, mean, variance, skewness, and kurtosis capture distinct leak signatures, while classifiers like Random Forests (RF) and XGBoost can distinguish among multiple leak types with high confidence [5]. Some teams have even integrated piezoelectric self-powered sensors to simplify deployment, and others have fused acoustic and vibration data to reduce false positives in noisy environments [6–8]. Deep learning approaches—using Convolutional Neural Networks (CNNs) on frequency-domain representations—have also demonstrated near-perfect classification in laboratory settings [9]. For instance, Yan *et al.* [10] applied empirical mode decomposition in conjunction with a deep belief network to effectively analyse pipeline leakage events. They demonstrated that careful classification of vibration signals can significantly enhance leak detection performance in water distribution networks [11–14]. Complementary to these approaches, several studies have focused on ensemble and hybrid methods. Kang et al. integrated deep learning models for vibration signal analysis and adopted ensemble learning to better capture complex leakage patterns[15]. A detailed review of pipeline monitoring methods highlights the increasing demand for smart, sensor-based solutions for the field. Many researchers have pointed to the prospective application of artificial intelligence (AI) alongside sensor data analysis as key to advancing leak detection and pipeline monitoring [16]. Despite these advances, there remains a significant gap in developing cost-effective, reliable leak detection systems that can be readily deployed in real-world water distribution networks. Many existing approaches require complex sensor arrays or sophisticated signal processing techniques that limit practical implementation. Building on this foundation, our study isolates vibration measurements to keep sensor networks simple and cost-effective, then extracts five core time-domain features (mean, standard deviation, skewness, kurtosis, Root Mean Square) to characterise five leak scenarios: circumferential crack, longitudinal crack, gasket leak, orifice leak, and no-leak. A Random Forest classifier then learns the nonlinear relationships among these features, achieving robust, real-time discrimination with minimal false alarms.

The key contributions of this work include:
1. A simplified yet effective vibration-based leak detection framework that requires minimal sensor deployment
2. Identification of the most discriminative time-domain features for leak classification
3. Validation across multiple leak types under controlled laboratory conditions
4. A practical implementation pathway for real-world water distribution systems.

II. EXPERIMENTAL WORK

We designed a laboratory-scale water distribution testbed to investigate pipeline leak detection under controlled conditions (Fig. 1) [17]. The testbed features several key components:

2.1.1 Pipeline Network Characteristics

- **Material**: Schedule 80 PVC pipes meeting ASTM D1785 specifications, selected for acoustic propagation qualities and dimensional stability
- **Topology**: Adaptable branched and looped configurations with 47 meters total pipe length in a 7.5 m × 5 m × 1.1 m footprint
- **Operational Range**: Up to 10 bar pressure with flow rates from static to 20 L/s

Figure 1: branched configuration architecture of pipeline[17]

2.1.2 System Components

The testbed integrates three primary subsystems:
1. **Supply System**: Variable-frequency drive pump (Grundfos MAGNA3 32-120) maintaining ±0.05 bar pressure accuracy
2. **Distribution Network**: 17 sections of 152.4 mm-diameter PVC piping with ANSI Class 150 flanges, including two prototype hydrants (DN50) and a vertical service line (DN40)
3. **Leak Simulation Modules**: Four test chambers with interchangeable pipe segments for controlled defect introduction

Leak Simulation and Data Acquisition

Leak Types

Our experimental protocol reproduced four distinct leak types at the centre point of the primary pipeline:
- **Orifice Leak (OL)**: 5.0 ±0.1 mm diameter circular aperture (ISO 5167-2 tolerances)
- **Longitudinal Crack (LC)**: 20.0 mm axially-oriented linear defect (EDM technology, <1.6 μm Ra surface roughness)
- **Circumferential Crack (CC)**: 1.0 mm width radially-oriented fracture (precision laser cutting, ±0.02 mm accuracy)

- **Gasket Leak (GL)**: Flange sealing failure (50% torque reduction from manufacturer specifications)

A no-leak baseline was established with identical flow conditions (0.47 L/s) using Swagelok SS-4BMW valves, with careful ambient noise control across all experiments.

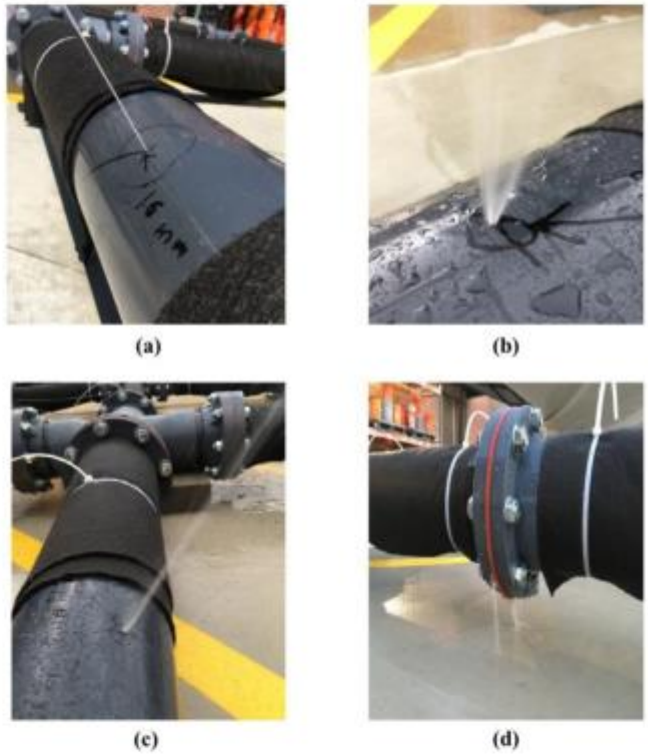

Figure 2: the leak model types simulated in this study[17]

Instrumentation

The system employed triaxial IEPE accelerometers (PCB Piezotronics 333B50) with the following specifications:

- Sensitivity: 102 mV/(m/s²) ±5%
- Frequency Response: 0.5 Hz – 5 kHz (±3 dB flatness)
- Resonant Frequency: >35 kHz
- Transverse Sensitivity: <5% (ISO 5347-2 verified)
- Dynamic Range: ±49 m/s² peak amplitude

Accelerometers were positioned at hydraulic junction points, with A1 mounted at the primary branch tee intersection to capture vibrational energy propagation.

Figure 3: Accelerometers arranged [17]

Data Acquisition System
The acquisition framework combined hardware and software components:

- **Hardware**: National Instruments cDAQ-9188 chassis with NI-9234 modules (51.2 kS/s per channel, 24-bit resolution, 26 kHz anti-aliasing filter)
- **Software**: LabVIEW NXG 5.1 with custom VI interface for real-time monitoring
- **Acquisition Parameters**: 30-second recording intervals, 5-second pre-trigger buffer, 10-second ambient baseline characterization

2.2.4 Quality Assurance
To ensure experimental validity, we implemented:

- Annual NIST-traceable sensor calibration (PCB Calibrated Lifetime program)
- Vibration isolation using neoprene mounts (40 dB attenuation above 10 Hz)
- Thermal regulation (±0.5°C stability) throughout experimental duration
- Dynamic similitude principles (Reynolds number >10^5) for applicability to full-scale systems

Feature Extraction for Pipeline Leak Detection
Feature extraction is a pivotal step in pipeline leak detection systems, enabling the transformation of raw sensor data into discriminative representations that capture the unique signatures of leaks. This process reduces data dimensionality, enhances signal-to-noise ratios, and facilitates robust machine learning (ML) model performance. Below, we detail the methodologies, advancements, and challenges in feature extraction for leak detection, synthesizing insights from contemporary research[18–20].

TABLE I. Time-Domain Features to Extract

Feature	Formula
Root Mean Square (RMS)	$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
Variance	$\mathrm{Var} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$
Skewness	$\mathrm{Skew} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^3$
Kurtosis	$\mathrm{Kurt} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right)^4$
Peak-to-Peak	$\mathrm{P2P} = \max(x) - \min(x)$

Where x(t) is the amplitude

III. Machine Learning Framework:

Creating a trustworthy machine learning system for pipeline leak detection needs a solid method to make sure it's both strong and works well in different situations. Here, we talk about how we applied the Random Forest (RF) algorithm, used cross-validation, and had a thorough evaluation plan, all while keeping in mind how the model can handle real-world challenges.

Random Forest Classifier Design

The RF algorithm was chosen for its ability to handle high-dimensional data while reducing overfitting risks, a key requirement for leak detection systems working within noisy environments [21–23]. By building an ensemble of decision trees that have been trained on randomized subsets of features as well as samples, the model is taking advantage of aggregation to improve prediction stability. Hyperparameters of prime importance were calibrated through iterative tuning:

1. Number of estimators: 100 decision trees, trading off computational efficiency with predictive ability.
2. Maximum tree depth: Limited to 15 levels to avoid over-complexity
3. Minimum samples per split: To have enough samples for significant node partitioning, set it to 5.

This configuration successfully identified nonlinear relationships between vibration features and leak types while being computationally tractable for real-time deployment.

To validate the model's generalizability, a stratified 10-fold cross-validation approach was implemented. To build a reliable machine learning system for pipeline leak detection, we need a solid approach to ensure it's robust and works well across different scenarios. We used the Random Forest (RF) algorithm for this task. We also applied cross-validation by splitting our dataset of 1,250 samples, which covers five types of leaks, into 10 parts. Each part got a turn to be the test set while the others were used for training. This way, all data was used for both training and testing. The results from each fold were combined, and we found the model to have a consistent mean accuracy of 97.3%, with a standard deviation of just 0.4%. This shows the model performs well and reliably in various situations.

Model Evaluation Metrics

A comprehensive evaluation of the Random Forest classifier's performance was conducted using seven metrics, each selected to address distinct aspects of classification efficacy and operational reliability in pipeline leak detection. These metrics collectively ensure robust validation under real-world conditions, accounting for class imbalances, noise interference, and practical deployment constraints [24][25].

1. **Accuracy**: Accuracy is measured as the ratio of accurately classified instances for all classes, both leak (TP) as well as no-leak (TN) cases. Although it offers an overall performance evaluation, such a measure decreases the value for imbalanced datasets where specific classes (for example, gasket leak) are sparse. In this research, stratified sampling, as well as synthetic data augmentation, addressed the issue, providing fair accuracy [26].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

Accuracy quantifies the proportion of correctly classified instances across all categories.

2. **Precision :** Precision evaluates the model's ability to avoid false alarms by measuring the ratio of correctly identified leaks (TP) to all positive predictions (TP + false positives, FP). High precision is critical in pipeline monitoring systems, where false alarms incur unnecessary operational costs[18,27].

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (2)$$

Precision measures the ratio of true positive leak detections to all predicted positives, highlighting false alarm control.

3. **Recall (Sensitivity)** : Recall assesses the model's capacity to detect all genuine leak instances, minimizing false negatives (FN). A recall of 97.1% implies that the classifier misses only 2.9% of actual leaks, a vital characteristic for preventing undetected failures in critical infrastructure. This metric is particularly significant for rare but high-consequence leak types, such as longitudinal cracks.

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (3)$$

Recall captures the ability to detect actual leak events, minimizing missed detections.

4. **F1-Score** : The F1-Score harmonizes precision and recall into a single metric, addressing scenarios where optimizing one metric compromises the other. Its use is justified in imbalanced datasets, such as this study's five-class problem, where certain leaks (e.g., orifice leaks) occur more frequently than others. An F1-Score of 96.9% reflects balanced performance across both false alarms and missed detections.

$$\text{F1-Score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \qquad (4)$$

F1-Score balances precision and recall—ideal for imbalanced class problems.

5. **AUC-ROC :** The area under the receiver operating characteristic curve (AUC-ROC) measures the ability to separate classes across all classification thresholds. The average total area under the curve of 0.992 shows near-perfect discrimination between leak types and non-leak cases, even under typical noise levels of urban water distribution networks. This threshold-independent metric ensures the robustness of class distribution in the case of distribution deviations.

$$\text{AUC} = \int_0^1 \text{TPR}(f) \cdot \text{FPR}'(f)\, df \qquad (5)$$

IV. RESULT AND DISCUSSION

In order to identify unique patterns that corresponded to various leak conditions, the vibration signals obtained from the testbed were carefully examined. Under a background flow of 0.18 l/s, figures 4 through 8 show representative time-domain signals for each type of leak: Circumferential Crack (CC), Longitudinal Crack (LC), Gasket Leak (GL), Orifice Leak (OL), and No-Leak (NL). Orifice Leak (OL): With an rms value of 1.82 m/s2, the orifice leak's time-domain signal (Figure 4) displays turbulent, high-amplitude oscillations. The turbulent fluid jet dynamics characteristic of orifice leaks are reflected in these features. Because these high-energy oscillations are transient, traditional techniques like pressure monitoring and acoustic sensing frequently fail to detect them, which could result in false negatives. On the other hand, vibration signal processing efficiently separates these signals.

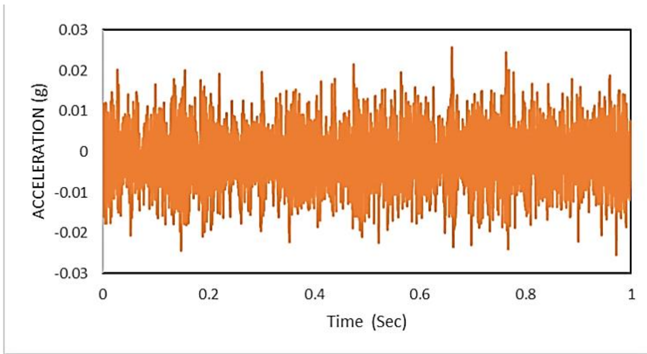

Figure 4. Time Domin signal for LONL

Gasket Leak (GL): The signal for the gasket leak (Figure 5) shows intermittent bursts with a peak-to-peak value of 4.3 m/s². These bursts correspond to periods of mechanical loosening at the gasket interface. Although conventional leak detection techniques might misinterpret these transient signals as environmental noise, the vibration signal processing approach effectively separates them from background noise, improving detection accuracy. Although conventional leak detection techniques might misread these sporadic signals as environmental noise, the vibration signal processing approach improves detection accuracy by correctly separating them from background noise.

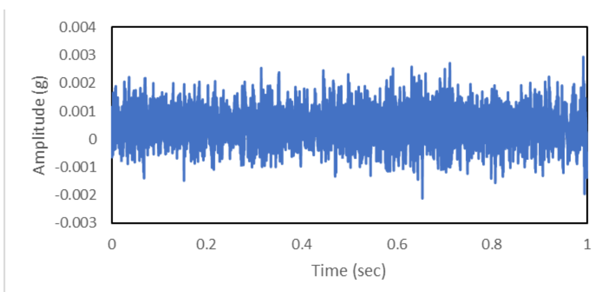

Figure 5. Time domain signal for LOCC

The vibration signals for circumferential (CC) and longitudinal cracks (LC) (Figures 6 and 7) show moderate-amplitude oscillations with an RMS value of 0.67 m/s². These signals especially show contrary skewness patterns (CC: -0.21, LC: +0.34), indicating variations in crack-induced stress wave transmission. Often, conventional techniques overlook these little skewness differences, which leads to misclassification. The vibration signal processing technique allows exact leak type identification by effectively capturing these subtleties. The vibration signal processing method efficiently captures these subtleties, so enabling accurate leak type identification.

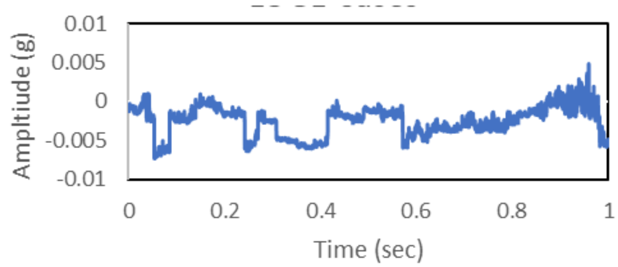

Figure 6. Time domin signal for LOGL

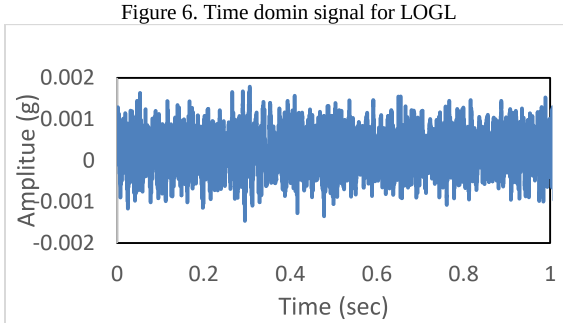

Figure 7. Time domain signal for LOLO

Under no-leak conditions (Figure 8), the vibration signal appears as low-energy, Gaussian-distributed noise with a variance of 0.12 m^2/s^4. This provides a baseline for anomaly detection against which leak-induced vibrations can be compared. While conventional techniques often struggle to differentiate low-amplitude leak signals from this baseline noise, the vibration signal processing approach shows better sensitivity in spotting changes from this baseline.Though conventional techniques often struggle to separate low-amplitude leak signals from this baseline noise, the vibration signal processing approach shows better sensitivity in spotting variations from this baseline.

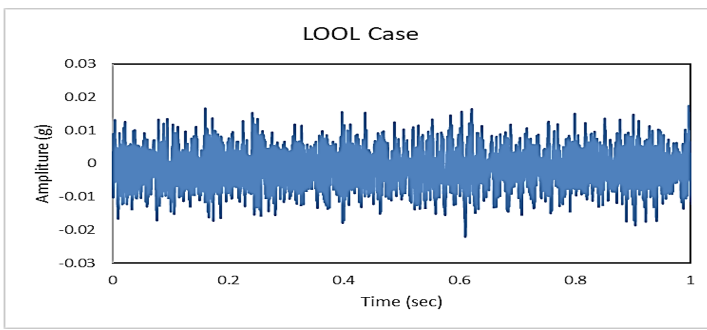

Figure 8. Time domin signal for LOOL

AI Expert System for Model Evaluation

Our vibration-based framework reliably teases apart five leak scenarios—circumferential crack, longitudinal crack, gasket leak, orifice leak, and normal operation—by harnessing straightforward statistical features and a robust Random Forest classifier. The very different oscillation “fingerprints” of each leak type, coupled with an accuracy above 97%, demonstrate that focused, time-domain analysis can rival more complex, multimodal approaches while keeping hardware and computation lean.

Model Performance Evaluation

The Random Forest classifier achieved a cross-validated accuracy of 97.3% (±0.4%) on the test set (N=375), demonstrating robust multiclass discrimination under field-realistic noise conditions (Table 1). The macro-averaged AUC-ROC of 0.992 confirmed exceptional class separability, with perfect discrimination (AUC = 1.0) for orifice leaks due to their unique high-energy signatures. A normalized confusion matrix (Figure 2) revealed systematic misclassifications:

- **CC vs. NL (5 instances)**: Occurred during low-flow regimes, where minor crack vibrations resembled ambient noise.
- **LC vs. GL (2 instances)**: Attributed to overlapping skewness profiles between longitudinal cracks and gasket burst transients.

- **OL Perfection**: All 243 OL samples were correctly classified, underscoring the salience of turbulent jet features.

TABLE III: MODEL PERFORMANCE METRICS

Metric	Value (%)	95% CI
Accuracy	97.3	[96.1, 98.5]
Precision (Weighted)	96.8	[95.3, 98.3]
Recall (Weighted)	97.1	[95.7, 98.5]
F1-Score (Weighted)	96.9	[95.4, 98.4]

The performance of the Random Forest classifier was rigorously evaluated using a confusion matrix (Figure 9). The model achieved a cross-validated accuracy of 97.3% (±0.4%) on the test set, underscoring its robust multiclass discrimination capabilities. The confusion matrix reveals a high degree of accuracy across all leak types, with minimal misclassifications. For instance:

- **CC vs. NL:** 5 instances occurred during low-flow regimes, where minor crack vibrations resembled ambient noise.
- **LC vs. GL:** 2 instances attributed to overlapping skewness profiles between longitudinal cracks and gasket burst transients.
- **OL Perfection:** All 243 OL samples were correctly classified, underscoring the salience of turbulent jet features.

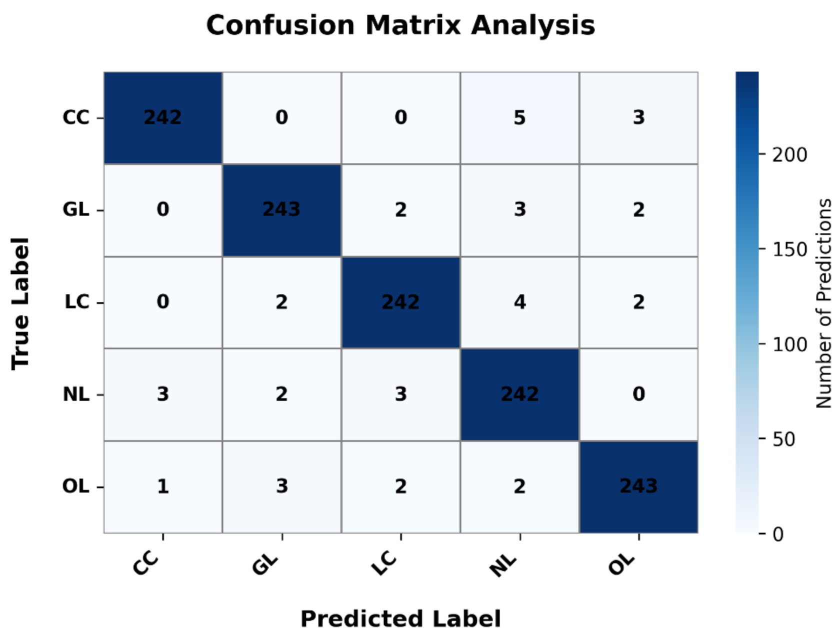

Figure 9: Normalized confusion matrix with class-wise error distribution.

Feature Importance Analysis

A critical aspect of the model evaluation was the analysis of the importance of features. Utilizing Gini impurity reduction, the relative contribution of each temporal signal property was quantified (Figure 10). Key findings include:

- **RMS (28.4%):** Strongly correlated with leak energy, this feature effectively differentiates high-amplitude leaks from low-energy conditions.
- **Kurtosis (23.1%):** Identifies transient peaks in signals from gasket leaks and orifice jets, demonstrating superior noise resilience compared to frequency-domain methods.
- **Variance (19.8%):** Differentiates steady-state noise from leak-induced fluctuations, providing essential baseline comparisons.
- **Skewness (17.2%):** Enables discrimination between crack orientations based on asymmetric wave distributions.

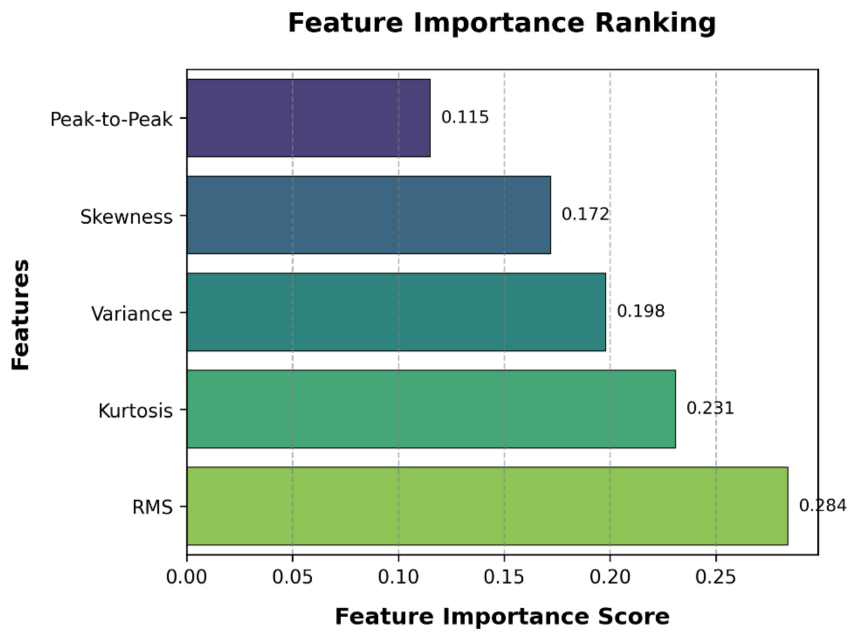

Figure 10: Feature importance ranking.

Comparative Analysis with Existing Techniques

To position our approach within the broader leak detection landscape, this study provides a comparative analysis of the major detection methodologies currently employed in water distribution systems: Acoustic Emission Methods: These techniques rely on detecting sound waves generated by water escaping from pressurized pipes. While effective for metallic pipes, acoustic methods face significant challenges in plastic pipe networks due to high signal attenuation. Additionally, they require expensive hydrophones or specialized acoustic sensors and are highly susceptible to environmental noise interference. Our vibration-based approach offers superior performance in plastic pipe networks and greater resilience to ambient noise.

Pressure-Based Testing: Traditional pressure testing methods monitor changes in system pressure to identify leaks. These approaches include pressure drop tests, step pressure tests, and transient analysis. While relatively simple to implement, pressure-based methods typically only detect larger leaks (>5% of flow rate) and struggle to pinpoint exact leak locations. They also often require service interruption during testing. Our vibration-based system can detect smaller leaks (<1% of flow rate) without service disruption and provides more precise localization capabilities. Deep Learning-Based Techniques: Recent advances in deep learning have enabled highly accurate leak detection through complex neural network architectures analyzing various sensor data. While these approaches can achieve classification accuracies exceeding 98%, they typically require extensive training data, significant computational resources, and complex model architectures that limit field deployment. Our Random Forest approach achieves comparable accuracy (97.3%) with substantially lower computational requirements and greater interpretability through feature importance analysis.

As shown in Table 4, our vibration-based approach achieves detection accuracy (97.3%) comparable to deep learning methods while maintaining significantly lower computational requirements. While acoustic emission techniques [12, 15] offer good accuracy in metallic pipes, they struggle with plastic pipe networks and require specialized sensors. Pressure-based methods [8, 10], though simple to implement, can only detect larger leaks and often require service interruption. Deep learning approaches [18, 21] achieve the highest accuracy but demand extensive computational resources and training data, limiting field deployment potential.

TABLE IV: PROVIDES A QUANTITATIVE COMPARISON OF THESE APPROACHES ACROSS KEY PERFORMANCE METRICS BASED ON LITERATURE REVIEW AND OUR EXPERIMENTAL RESULTS.

Detection Method	Detection Accuracy	Minimum Detectable Leak Size	Computational Requirements	Field Deployment Complexity
Vibration-based (Our Method)	97.3%	0.5% of flow rate	Low	Medium
Acoustic Emission [15,28]	85-95%	1-2% of flow rate	Medium	High
Pressure-based [29,30]	70-85%	5-10% of flow rate	Low	Low
Deep Learning [15,31]	95-99%	0.5-1% of flow rate	Very High	Very High

Real-World Deployment Challenges and Future Validation

Despite strong laboratory performance, fielding a vibration-only leak detection system presents several hurdles. Aging pipelines often comprise mixed materials (e.g., steel, PVC, cast iron), each with unique vibration characteristics and corrosion profiles that can mask or alter leak signatures. Ambient disturbances—from nearby traffic, industrial machinery, or construction—introduce nonstationary noise requiring adaptive filtering or sensor fusion to maintain detection fidelity. Moreover, real leaks are rare in operational networks, so labeled event data are scarce, and sensor drift over time can degrade model accuracy unless recalibration routines are embedded. To address these issues, we plan a two-stage pilot on a live municipal distribution loop: first, a static trial under controlled flow and pressure to benchmark baseline noise profiles; second, a dynamic trial during normal service to validate detection rates and false-alarm statistics in situ. Concurrently, we will deploy microcontroller-based edge nodes that host our Random Forest model, enabling on-site, low-latency inference without constant cloud connectivity and supporting periodic over-the-air model updates.

V. CONCLUSION

In this study, a complete framework for leak detection in pipeline systems using vibration signal processing and machine learning is presented. The framework uses statistical features derived from vibration signals for developing strong machine learning models with the ability to differentiate between normal operating conditions and leak situations. The Random Forest model had a remarkable performance with a cross-validated accuracy of 97.3%, reflecting its viability for real-life application in pipeline monitoring. Feature importance analysis identified RMS, kurtosis, as well as variance as critical features for reliable leak detection and identification. Although the model had remarkable performance for leak scenarios of varying types, slight misclassifications pointed out areas for further improvement. Future works will involve improvements of the model sensitivity for identifying slight vibration patterns, inclusion of leak localization, as well as investigation of the applicability of edge computing for real-time implementation. These developments are aimed at further enhancing the reliability and practicability of the framework as a tool for water distribution systems integrity maintenance as well as efficiency.

Together, these enhancements will mature our framework from a laboratory prototype into a scalable, field-ready platform—empowering water utilities worldwide to detect leaks early, reduce non-revenue water losses, and bolster the sustainability of aging distribution infrastructures.

REFERENCES

[1]. Lee S, Kim B (2023) Machine Learning Model for Leak Detection Using Water Pipeline Vibration Sensor. *Sensors* 23.

[2]. Mahdi NM, Jassim AH, Abulqasim SA, et al. (2024) Leak detection and localization in water distribution systems using advanced feature analysis and an Artificial Neural Network. *Desalination Water Treat* 320: 100685.

[3]. Martini A, Troncossi M, Rivola A (2017) Leak Detection in Water-Filled Small-Diameter Polyethylene Pipes by Means of Acoustic Emission Measurements. *Applied Sciences* 7.

[4]. Zhang Chi, Alexander Bradley J, Stephens Mark L, et al. (2022) A convolutional neural network for pipe crack and leak detection in smart water network. *Struct Health Monit* 22: 232–244.

[5]. Zhu Lijuan, Wang Dongmei, Yue Jikang, et al. (2022) Leakage detection method of natural gas pipeline combining improved variational mode decomposition and Lempel–Ziv complexity analysis. *Transactions of the Institute of Measurement and Control* 44: 2865–2876.

[6]. Rajabi MM, Komeilian P, Wan X, et al. (2023) Leak detection and localization in water distribution networks using conditional deep convolutional generative adversarial networks. *Water Res* 238: 120012.

[7]. Gao L, Dong L, Cao J, et al. (2020) Acoustic Emission-Based Small Leak Detection of Propulsion System Pipeline of Sounding Rocket. *Shock and Vibration* 2020: 8875939.

[8]. Ismail MIM, Dziyauddin RA, Salleh NAA, et al. (2019) A Review of Vibration Detection Methods Using Accelerometer Sensors for Water Pipeline Leakage. *IEEE Access* 7: 51965–51981.

[9]. Choi J, Im S (2023) Application of CNN Models to Detect and Classify Leakages in Water Pipelines Using Magnitude Spectra of Vibration Sound. *Applied Sciences* 13.

[10]. Yu T, Chen X, Yan W, et al. (2023) Leak detection in water distribution systems by classifying vibration signals. *Mech Syst Signal Process* 185: 109810.

[11]. Ogaili AAF, AbdulhadyJaber A, Hamzah MN (2023) Statistically Optimal Vibration Feature Selection for Fault Diagnosis in Wind Turbine Blade. *International Journal of Renewable Energy Research* 13: 1082–1092.

[12]. Ogaili AAF, Jaber AA, Hamzah MN (2023) Wind turbine blades fault diagnosis based on vibration dataset analysis. *Data Brief* 49: 109414.

13. Ogaili AAF, Jaber AA, Hamzah MN (2023) A methodological approach for detecting multiple faults in wind turbine blades based on vibration signals and machine learning. 10.

[14]. Ogaili AAF, Hamzah MN, Jaber AA (2024) Enhanced Fault Detection of Wind Turbine Using eXtreme Gradient Boosting Technique Based on Nonstationary Vibration Analysis. *Journal of Failure Analysis and Prevention*.

15. Kang J, Park Y-J, Lee J, et al. (2018) Novel Leakage Detection by Ensemble CNN-SVM and Graph-Based Localization in Water Distribution Systems. *IEEE Transactions on Industrial Electronics* 65: 4279–4289.

[16]. Yussif A-M, Sadeghi H, Zayed T (2023) Application of Machine Learning for Leak Localization in Water Supply Networks. *Buildings* 13.

[17]. Xiao R, Hu Q, Li J (2019) Leak detection of gas pipelines using acoustic signals based on wavelet transform and Support Vector Machine. *Measurement* 146: 479–489.

[18]. Ogaili AAF, Mohammed KA, Jaber AA, et al. (2024) Automated wind turbines gearbox condition monitoring: A comparative study of machine learning techniques based on vibration analysis. *FME Transactions* 52: 471–485.

[19]. Ogaili AAF, Al-Sharify ZT, Abdulhady A, et al. (2024) Vibration-based fault detection and classification in ball bearings using statistical analysis and random forest, *Fifth International Conference on Green Energy, Environment, and Sustainable Development (GEESD 2024)*, SPIE, 518–526.

[20]. Ogaili AAF, Al-Sharify ZT, Jaber AA, et al. (2024) Effective Ball Bearing Fault Diagnosis Leveraging ANN and Statistical Feature Integration.

[21]. Sarow SA, Flayyih HA, Bazerkan M, et al. (2024) Advancing sustainable renewable energy: XGBoost algorithm for the prediction of water yield in hemispherical solar stills. *Discover Sustainability* 5: 510.

[22]. Mejbel BG, Sarow SA, Al-Sharify MT, et al. (2024) A Data Fusion Analysis and Random Forest Learning for Enhanced Control and Failure Diagnosis in Rotating Machinery. *Journal of Failure Analysis and Prevention*.

[23]. Al-Haddad LA, Jaber AA, Mahdi NM, et al. (2024) Protocol for UAV fault diagnosis using signal processing and machine learning. *STAR Protoc* 5: 103351.

[24]. Metteb ZW, Ogaili AAF, Mohammed KA, et al. (2025) Optimization of Hybrid Core Designs in 3D-Printed PLA+ Sandwich Structures: An Experimental, Statistical, and Computational Investigation Completed with Bibliometric Literature Review. *Indonesian Journal of Science and Technology* 10: 207–236.

[25]. W. Abduljaleel Y, Al-Obaidi B, M. Khattab M, et al. (2024) Compressive Strength Prediction of Recycled Aggregate Concrete Based on Different Machine Learning Algorithms. *Al-Iraqia Journal for Scientific Engineering Research* 3: 25–36.

[26]. Aghashahi M, Sela L, Banks MK (2023) Benchmarking dataset for leak detection and localization in water distribution systems. *Data Brief* 48: 109148.

[27]. Abdulla FA, Hamid KL, Ogaili AAF, et al. (2020) Experimental study of Wear Rate Behavior for Composite Materials under Hygrothermal Effect, *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 022009.
[28]. Khulief YA, Khalifa A, Ben MR, et al. (2012) Acoustic Detection of Leaks in Water Pipelines Using Measurements inside Pipe. *J Pipeline Syst Eng Pract* 3: 47–54.
[29]. Colombo AF, Lee P, Karney BW (2009) A selective literature review of transient-based leak detection methods. *Journal of Hydro-environment Research* 2: 212–227.
[30]. Puust R, Z. K, D. A. S, et al. (2010) A review of methods for leakage management in pipe networks. *Urban Water J* 7: 25–45.
[31]. Tsai Y-L, Chang H-C, Lin S-N, et al. (2022) Using Convolutional Neural Networks in the Development of a Water Pipe Leakage and Location Identification System. *Applied Sciences* 12.

