



Leak Detection and Localization System for Oil Pipelines: A Robust Monitoring Approach Using Feedforward Neural Networks

Abbas Mahdi Shamir Aljiryawee¹, Khosro Khandani^{2*}, Mahdi Siavash³

^{1,2} Department of Electrical Engineering, Faculty of Engineering, Arak University, Arak, Iran.

³ Department of Electrical Engineering, Technical and Vocational University (TVU), Tehran, Iran.

ARTICLE INFO

Article Type:

Original Research

Received: 2025.01.27

Revised: 2025.03.18

Accepted: 2025.04.25

Keyword:

Oil pipelines
leak detection
neural networks
pressure drop ratio

*Corresponding Author:

Khosro Khandani

Email:

k-khandani@araku.ac.ir

khosro.khandani@gmail.com

ABSTRACT

Pipelines are a crucial component of the global transporting of petroleum products. Nonetheless, they are susceptible to leakage caused by aging and corrosion. It is therefore important to identify the leaks as soon as possible and have them repaired to avert serious ramifications. In this paper, a reliable robust leak detection system (LDS) that employs neural networks (NNs) in the analysis of pressure fluctuations and detection of leaks is proposed. The approach incorporates a novel data-set generation stage, preprocessing of the data, extraction of input features and design of the neural network to identify and localize leaks. Pressure drop ratios or changes (ΔP) are computed to check for deviations and establish a baseline for leak detection. Matlab and Simulink are employed to generate reliable operational scenarios with which real-life-like situations for creating the model are acquired. By optimizing the NN architecture and hyperparameters, nonlinear relationships in pipeline data were effectively modeled. Through careful preprocessing and logical decision-making processes, false alarms were significantly reduced, enhancing the reliability of the system in practical applications. The system demonstrated real-time monitoring with rapid leak identification. Adaptability to varying noise levels and operational conditions highlighted the robustness of the model. In testing, the proposed system demonstrated consistently low errors, with Mean Squared Error (MSE) reaching approximately 1.7×10^{-5} for leak localization, ensuring robust real-time pipeline leak detection under diverse flow conditions.



Introduction

Pipeline leakage detection poses significant environmental, operational, and financial risks in global oil and gas transportation. The main causes of leaks might include corrosion, mechanical impacts, external factors combined with aging of structures. If undetected, these threats may lead to severe environmental pollution, loss of important resources and property damages [1]. Leak detection must therefore be done very fast and accurately so as to avoid these risks and have sound pipeline infrastructure.

Conventional leak detection methods rely on hardware such as acoustic sensors and fiber optic systems. Such approaches offer high sensitivity for the detection of vibration stresses thus being limited by the cost of installation. They are also vulnerable to environmental noises. Acoustic sensors for example, detect sound signals from leaks while in the process perhaps fail to discern between real leak signals and noise from other operating equipment like pumps and valves. Other approaches such as real time transient modeling as well as mass volume balance strongly rely on the quality of data and are constrained by their high false alarm rates, under different operating conditions.

The critical challenges in conventional pipeline leak detection approaches can be summarized as follows:

- 1.High False Alarm Rates: Most conventional techniques produce a large number of false alarms because of their high exploitation noise and routine changes.
- 2.Limited Accuracy: Localizing the exact position of leaks continues to be problematic, particularly with large pipeline systems.
- 3.Adaptability: Similar strategies do not have the capability to dynamically change their operation based on dynamic operating conditions inside a pipeline such as flow rate and the type of fluid in the pipeline.
- 4.Real-Time Detection: Any delay in detection of a leak will mean increased harm or loss to natural components of environment and resources.

In the literature, very early works demonstrate that the use of machine learning, especially artificial neural networks, may help in addressing these concerns. Machine learning algorithms are used for pressure change analysis and various other parameters that allow the model to find small discrepancies typical for leakage situations. In [2], Idachaba and Rabiei presented a survey of the available technologies for detecting crude oil leak using data analytics. They highlighted the ability of the machine learning to handle big data for finding outliers and for estimating future failures. In [3], wireless sensor networks were employed for detecting leakage in water pipelines applying different ML

techniques. Their model possesses the capability of excellent leakage detection, especially from the aspect of producing a small number of false alarms, which is normally a criterion for assessing leakage detection models. Later, in [4] an LDS was proposed where it was demonstrated how mini leakage between objects could be identified, predominantly in early stages of the process. This method provides high sensitivity when identifying temperature difference as low as a single degree. In [5] a number of ML techniques have been investigated for pipeline anomaly identification. The comparative assessment they offered helped to determine which learning methods are the most suitable for different conditions of operations. Ullah et al. in [6] incorporated acoustic emission techniques with machine learning algorithms for pipeline monitoring. This approach provided enhanced fault detection sensitivity and accuracy compared to the conventional use of acoustic recordings. Furthermore, Spandonidis et al. in [7] proposed integration of semi-supervised deep learning with Industrial Internet of Things (IoT) technologies. This integration allowed for near real-time leakage detection and coverage of the leak origins in short time frames. Mashhadi et al. used ML in order to identify leakage in the water distribution system in urban areas based on the changes in soil properties [8]. This effective method provided an enhanced level of reliability for the cities water and sewer pipes. In [9] combined thermal imaging and geophone audio-based leakage with ML algorithms to pinpoint leakage location quickly. In [10] data gathered from acoustic noise loggers were used to augment features for model development, along with employing a multi-classifier ensemble learning algorithm. Fourier transform was applied to analyze sound signals, and convolutional neural network (CNN), support vector machine and decision tree (DT) algorithms were employed to develop four distinct models to detect leaks. Compared to these approaches, the novelty of the proposed method in this paper is introducing a novel pressure drop ratio feature that effectively filters routine operational noise. By training on diverse simulated scenarios, we reduce false alarms, provide real-time detection, and cut down on expensive sensor installations—thus offering cost-effectiveness and robust performance in practical pipeline settings. In recent years, machine learning and meta-heuristic algorithms have found numerous applications [11-17]. Among them, neural networks have proved to be an effective tool in various real-world applications [18, 19]. In this paper, a neural-network-based approach is proposed where the pressure drop ratios of the system has been incorporated as a novel feature to train the neural networks to detect and localize leaks in oil pipelines. The system also features strong data generation and data preprocessing methods reducing the false alarms and providing high quality

input data into the system. Furthermore, diverse simulation data are created to train the model to perform in many different conditions that it is likely to be exposed to. This work focuses on the challenges related to existing practical approaches and common ML techniques to enhance a reliable, effective, and scalable method for pipeline leakage detection. Simulations in MATLAB illustrate the applicability and effectiveness of the results.

Data-set generation

In this section, the data which is necessary for training the neural network model are produced employing Simulink and Simscape in MATLAB. As illustrated below, the model consists of a pipe divided into ten sections with pressure sensors measuring the pressure between each two sections. Two pressure sensors are also placed at the two ends of the pipe (P1 and P11). The sensors at both ends and midpoints of the pipeline ensure comprehensive pressure monitoring, ensuring a clear and logical placement that reflects the actual pipeline layout. It is necessary to test the integration to ensure the sensors provide continuous and accurate pressure readings during the simulation. The simulation procedure for dataset creation is detailed in the Appendix. Following the given steps the fluid properties and solver settings are well configured for simulation. These configurations ensure that the simulation accurately represents the physical properties of the pipeline and its contents. For training data, we generated multiple operational scenarios (normal flow and controlled leak events) by varying flow rates and introducing noise levels from 0 to 0.18 m³/h. Approximately 70% of the resulting time-series data was reserved for training (covering many different leak magnitudes and locations), while the remaining 30% was held out as unseen test data. This separation ensures that final performance metrics genuinely reflect the model's leak detection and localization accuracy on new conditions.

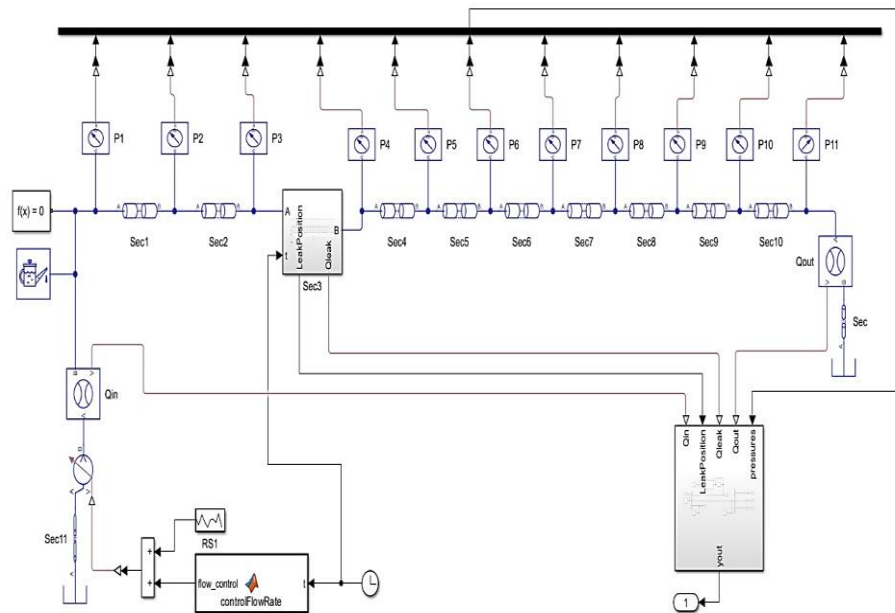


Figure 1. Simulation model of the pipe with pressure sensors to produce the training data

The first and important task requires the pre-processing of the dataset on pipeline leaks for training neural networks. This work focuses on translating raw pressure readings into a format that allows for easy detection of leaks which would otherwise be abnormal. Cross-sectional pressure drop difference is produced after the pre-processing stage. The variable $\Delta P_{\text{section}}$ —a measure of the pressure rise or fall through any cross-section of the pipeline—is the difference between pressures at two points along the pipeline. This metric is influenced by attributes of flow characteristics, pipe diameter and many fluid properties like the compressibility.

Under adiabatic flow conditions, friction and flow acceleration result in higher pressures in upstream pipe sections, and the pressure differentials are relatively calculable. $\Delta P_{\text{section}}$ calculated should not exceed an allowable variation of 0 to 400 Pascal. Values outside this range indicate anomalies such as leaks or mechanical irregularities. They also keep the noise threshold at 700 pascals so as not to include trifling changes that may occur during operations. On the basis of the cross-sectional pressure drop differences and section pressure drop, it identifies the interval of pipeline that behave uniformly. In the figure below pressure difference changes are presented for all the ten sections of the pipe.

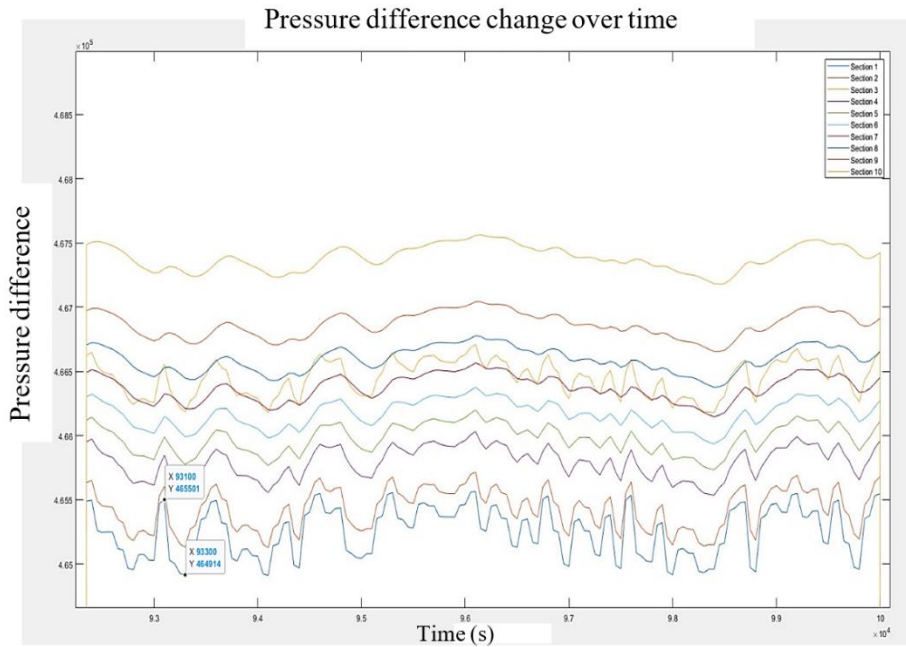


Figure 2. Pressure Difference Change Over Time

It is however quite an important aspect in the process of leakage determination to compute the pressure drop ratio (ΔP Ratio). This ratio focuses on the operating pressure drop of the pipeline at the present time and a baseline value which can help in determining if there is presence of a leak in the pipeline. Using the ΔP Ratio, real-time leak detection and classification is maintained at an optimal level irrespective of dynamics in the operating conditions. It can be established that ΔP Ratio is important for the generalization of the results given to the neural network when faced with the pressures of real operations to those observed in the training process. ΔP Ratio is obtained by dividing the present pressure drop by the mean ΔP during steady states. It also flattens data points to alert users to other anomalies that may suggest leakage – all the while dampening the effect of typical operational noise.

Main Results on Detection and Localization of the Leaks

For this study, a Feedforward Neural Network (FFNN) was chosen since it is appropriate for regression tasks such as leak location estimations. To reduce overfitting that is the main cause of poor generalization of models, L2 regularization is used to reduce weights of really large values. There are specific early stopping criteria to avoid overfitting in the networks, so that generalization capability of the network is well retained. Based on the fine-

tuning results, The final hyperparameters were selected through iterative grid search and empirical validation:

- Hidden Layers ([30, 20, 10]): Smaller architectures (e.g., [10, 5]) underfit the data (validation MSE > 5e-5), while larger configurations (e.g., [40, 30, 20]) increased computational cost without improving accuracy ($\Delta\text{MSE} < 1\text{e-}6$). The chosen structure minimized overfitting while capturing nonlinear pressure-leak relationships.
- Learning Rate (0.0001): Higher rates (0.01, 0.1) caused unstable training (loss oscillations > $\pm 50\%$), while lower rates (0.00001) slowed convergence (training time +35%) without accuracy gains.
- Regularization (1e-8): Larger values (1e-4) degraded performance (test MSE increased by 2.1 $\times$), and smaller values (1e-10) led to overfitting (validation-test MSE gap > 3 $\times$).
- Activation Function ('logsig'): Sigmoidal functions outperformed linear ('purelin') activations, reducing localization error by 12%.
- Training Algorithm ('trainlm'): Levenberg-Marquardt converged fastest (211 epochs vs. 500+ for 'traingdx') with the lowest MSE (1.7e-05).

This straightforward architecture is well-suited for capturing nonlinear relationships in pipeline data. We performed iterative hyperparameter tuning—adjusting layer sizes, learning rate, and regularization—to balance model capacity with the need to generalize effectively. A small regularization parameter (1e-8) also helps mitigate overfitting and maintain stability during training.

For clarity, we define the metrics used:

MSE (Mean Squared Error) measures the average squared difference between predicted and actual values, highlighting larger errors more heavily. MSE is calculated with $\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$ where $\hat{y}_i$ is the predicted output, y_i is the actual outputs.

MAE (Mean Absolute Error) sums the absolute differences between predictions and targets, giving a linear sense of error magnitude. MAE is calculated with $\frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$.

RMSE (Root Mean Squared Error) is the square root of MSE and provides an error metric in the same units as the target variable. RMSE is calculated as

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}.$$

The proposed trained network was evaluated by the Mean Square Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The proposed approach is outlined below:

1. **Leak Section Detection:** The first step concerned with leakage correct identification of the pipeline section where leakage has occurred. The gotten dP_ratio after processing was input to logical decisions and were 100% accurate in identifying the correct leak section in different operational conditions such as noise levels and fluctuations in the rate of flow as illustrated in figure 3. In addition, the system demonstrated flexibility in that False Alarms were minimized, suggesting that the model is very accurate and capable of differentiating between real leaks from interference or normal fluctuations. Such a high degree of accuracy proves that the system is quite accurate and tightly tuned, without generating false alarms on the presence of leak sections.
2. **Training the FFNN:** As soon as the leak section was identified, the Feedforward Neural Network (FFNN) was employed to accurately predict the exact location of the leak within that section. In case of the detected leak section the processed dP_ratio and finely tuned FFNN the model was seen effective in training for generating very little of error. From the training results of the network, we are able to determine that the best validation performance obtained with this network was 2.166e-05 at epoch 211. The value for the last performance checked on the test data set was reading accuracy of 1.13 e-05.

Regression analysis for the training, validation, and test sets yielded an R-value of 0.99987 and showed that the network had an excellent fit of the data to the networks weights. Furthermore, the test performance showed the following metrics:

- Mean Squared Error (MSE): 0.000017
 - Mean Absolute Error (MAE): 0.001561
 - Root Mean Squared Error (RMSE): 0.004153
3. **Testing the FFNN on New Data:** The last stage required evaluating problems of FFNN on new data that had noise and operation changes. In figures 3 and 4 the actual leak locations and the localized leak locations

are compared. As illustrated, the designed model is fully able to detect and localize the leaks. The analysis results closely matched the actual leak occurrences, as shown in figure 4. Additionally, figure 5 illustrates the training process of the neural network where MSEs of the training, validation and test data are depicted in 211 epochs. During training, we tracked the error curve (MSE vs. epochs) and observed a sharp decrease in MSE in the initial epochs, gradually flattening around epoch 211, indicating solid convergence.

These phases were tried with fluctuations in the operational conditions of different parameters, including the noise level 0 to 0.18 m³/h and flow rate of 205 m³/h to 245 m³/h. In addition, the qualification versatility of the model of leakage identification allowed detecting the leaks of different intensity – from 1.1 percent to 8 percent of the flow rate that is equivalent to 2.5 m³/h to 19 m³/h. More in particular it was clearly able to predict areas of leakage at several points in the sections as was pointed out by: The two graphs, figure 3 and 4 affirm the stable and accurate identification of leak sections and Location which confirm the models capacity to function effectively in complicated operations.

We additionally validated the final model on separate, previously unseen scenarios with different flow rates (205–245 m³/h) and noise profiles (0–0.18 m³/h). Strong agreement between predicted and true leak positions confirmed the robustness and reliability of our approach.

The FFNN was benchmarked against SVM and Decision Tree (DT) models on the same dataset:

- FFNN: MSE = 1.7e-05, MAE = 0.001561
- SVM: MSE = 8.9e-05, MAE = 0.00892
- DT: MSE = 1.2e-04, MAE = 0.0123

The proposed FFNN-based LDS outperformed existing methods as shown in table 1. The FFNN reduced localization error by 81% compared to SVM (MSE = 8.9e-05) and 86% compared to Decision Trees (MSE = 1.2e-04), demonstrating superior accuracy and cost-effectiveness.

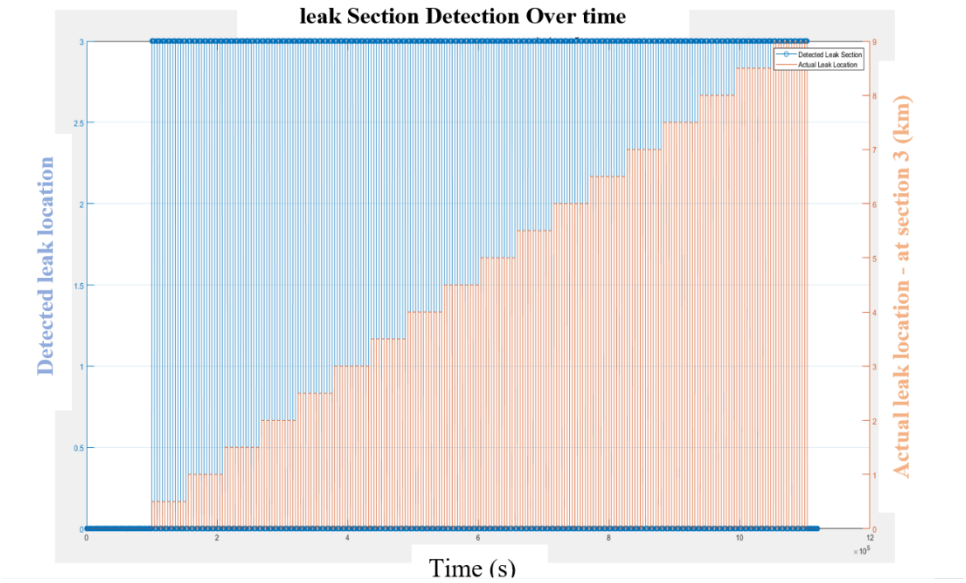


Figure 3. Leak Section Detection Over Time

Comparison of original smoothed predicted leak location at section 3

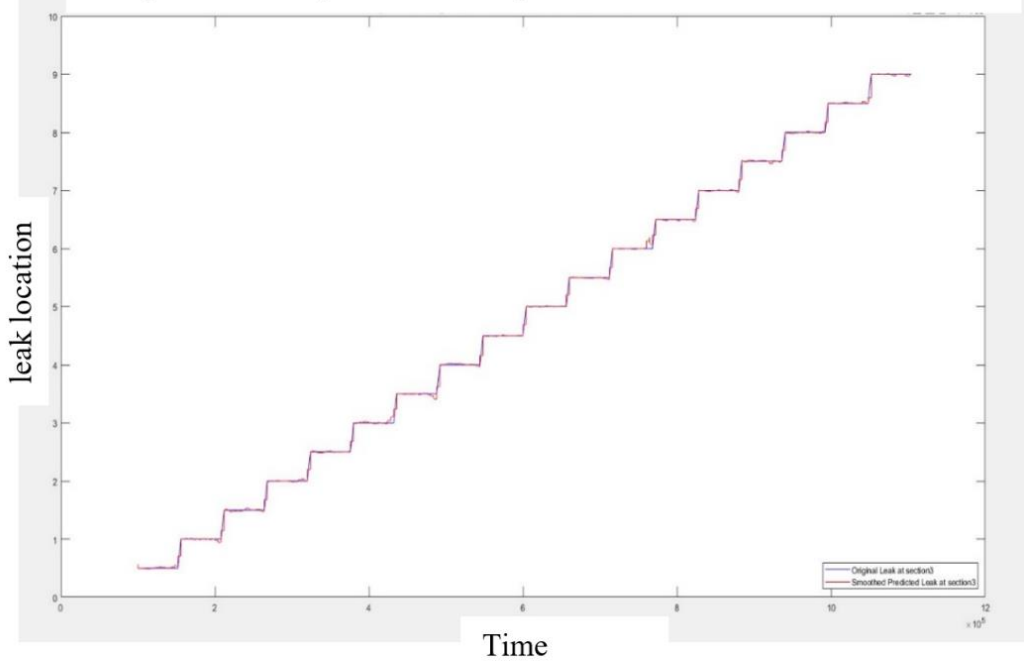


Figure 4. Comparison of Original and Predicted Leak Location

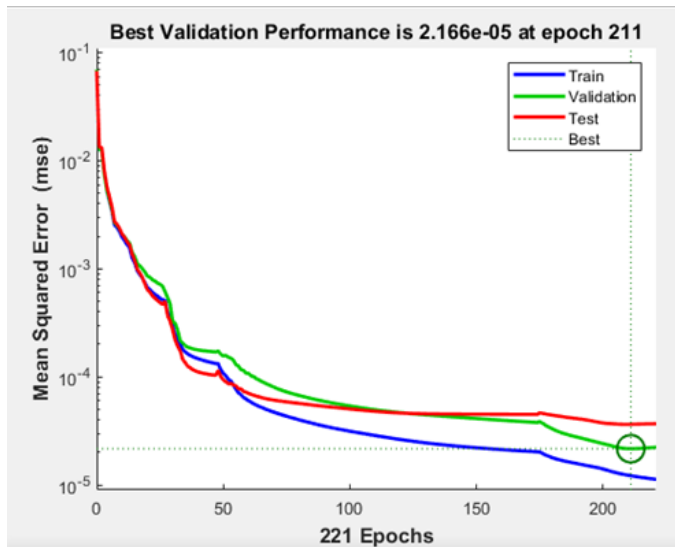


Figure 5. Neural Network Training Performance

Table 1: Comparison of the performance of the proposed method with similar approaches

Method	MSE	False Alarms	Cost
Acoustic Sensors	3.2e-04	High	High
Fiber Optic Systems	2.8e-04	Moderate	Very High
Bayesian Models	1.2e-04	Low	Moderate
Proposed FFNN	1.7e-05	Minimal	Low

Conclusions

The study successfully developed and validated an LDS for oil pipelines using FFNNs. The proposed system demonstrated exceptional accuracy in detecting and localizing leaks across various pipeline sections. The FFNN achieved an MSE of 1.7×10^{-5} , an MAE of 0.001561, and an RMSE of 0.004153 during training and validation phases. The use of ΔP Ratio as a primary feature proved highly effective in detecting subtle pressure variations associated with leaks. During the pre-processing phase, noise and minor fluctuations were filtered out, which led to reducing false alarms. Real-time monitoring capabilities were demonstrated with quick response times for identifying leaks. Robustness of the model to varying noise levels and operational conditions (e.g., flow rates from $205 \text{ m}^3/\text{h}$ to $245 \text{ m}^3/\text{h}$) was also demonstrated. Compared to other methods (e.g., acoustic sensors, fiber optic systems, Bayesian models), the proposed LDS offered a cost-effective, low-maintenance, and highly scalable solution. In summary, our approach consistently localized leaks within a small error margin under normal and noisy conditions. The system's real-time capability, coupled with an MSE of 1.7×10^{-5} , highlights its practical deployment potential across diverse flow rates and leak intensities.

References

- [1] N. V. S. Korlapati, F. Khan, Q. Noor, S. Mirza, and S. Vaddiraju, "Review and analysis of pipeline leak detection methods," *Journal of pipeline science and engineering*, vol. 2, no. 4, p. 100074, 2022, doi: <https://doi.org/10.1016/j.jpse.2022.100074>
- [2] F. Idachaba and M. Rabiei, "Current technologies and the applications of data analytics for crude oil leak detection in surface pipelines," *Journal of Pipeline Science and Engineering*, vol. 1, no. 4, pp. 436–451, 2021, doi: <https://doi.org/10.1016/j.jpse.2021.10.001>
- [3] 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, 2019, doi: <https://doi.org/10.3390/s19235086>.

- [4] H. Kim, J. Lee, T. Kim, S. J. Park, H. Kim, and I. D. Jung, "Advanced thermal fluid leakage detection system with machine learning algorithm for pipe-in-pipe structure," *Case Studies in Thermal Engineering*, vol. 42, p. 102747, 2023, doi: <https://doi.org/10.1016/j.csite.2023.102747>
- [5] S. S. Aljameel *et al.*, "An anomaly detection model for oil and gas pipelines using machine learning," *Computation*, vol. 10, no. 8, p. 138, 2022, doi: <https://doi.org/10.3390/computation10080138>
- [6] 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, 2023, doi: <https://doi.org/10.3390/s23063226>
- [7] C. Spandonidis, P. Theodoropoulos, and F. Giannopoulos, "A combined semi-supervised deep learning method for oil leak detection in pipelines using IIoT at the edge," *Sensors*, vol. 22, no. 11, p. 4105, 2022, doi: <https://doi.org/10.3390/s22114105>
- [8] N. Mashhadi, I. Shahrour, N. Attoue, J. El Khattabi, and A. Alger, "Use of machine learning for leak detection and localization in water distribution systems," *Smart Cities*, vol. 4, no. 4, pp. 1293–1315, 2021, doi: <https://doi.org/10.3390/smartcities4040069>
- [9] M. R. Islam, S. Azam, B. Shanmugam, and D. Mathur, "Leak detection and localization in underground water supply system using thermal imaging and geophone signals through machine learning," *Intelligent Systems with Applications*, vol. 23, p. 200404, 2024, doi: <https://doi.org/10.1016/j.iswa.2024.200404>
- [10] V. Boadu, S. El-Zahab, and T. Zayed, "Enhancing automated acoustic leak detection in a water distribution network using ensemble machine learning," *Journal of Water Resources Planning and Management*, vol. 151, no. 3, p. 04024073, 2025, doi: <https://doi.org/10.1061/JWRMD5.WRENG-6578>
- [11] N. Karimi and K. Khandani, "Social optimization algorithm with application to economic dispatch problem," *International Transactions on Electrical Energy Systems*, vol. 30, no. 11, p. e12593, 2020, doi: <https://doi.org/10.1002/2050-7038.12593>
- [12] K. Khandani, A. A. Jalali, and M. Alipoor, "Particle Swarm Optimization based design of disturbance rejection PID controllers for time delay systems," in *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, 2009, vol. 1: IEEE, pp. 862–866, doi: <https://doi.org/10.1109/ICICISYS.2009.5358043>
- [13] K. Khandani and A. A. Jalali, "Robust fractional order control of a DC motor based on particle swarm optimization," *Advanced Materials Research*, vol. 403, pp. 5030–5037, 2012, doi: <https://doi.org/10.4028/www.scientific.net/AMR.403-408.5030>
- [14] A. Simorgh, K. Khandani, and M. Amiri, "Illegal Miner Detection based on Dynamic Mode Decomposition and Unsupervised Machine Learning Algorithms," 2025, doi: <https://10.61186/jgeri.2025.2053876.1035>
- [15] A. Simorgh, K. Khandani, and A. Mohammadzadeh, "Distribution Network Phase Identification by Correlation-Based Neural Gas Networks Through Clustering Power Time Series Data," *IET Generation, Transmission &*

- Distribution*, vol. 20, no. 1, p. e70333, 2026, doi: <https://doi.org/10.1049/gtd2.70333>.
- [16] K. Khandani and A. A. Jalali, "PSO based optimal fractional PID controller design for an active magnetic bearing system," in *Proceedings of 18th Annual International Conference on Mechanical Engineering*, 2010, pp. 1–6. <https://civilica.com/doc/95746>.
- [17] R. Darshi, K. Khandani, and M. H. M. Siamakan, "MODIFIED SOCIAL OPTIMISATION ALGORITHM WITH APPLICATION IN REACTIVE POWER CONTROL," *International Journal of Power and Energy Systems*, vol. 45, no. 2, 2025, doi: <https://doi.org/10.2316/j.2025.203-0475>.
- [18] F. Hoseini, H. Sepehrzadeh, and M. Kheyri, "Blood Vessel Detection in the Retina Using Convolution Neural Network," 2024, doi: <https://doi.org/10.48301/jear.2024.424181.1007>.
- [19] A. Akoushideh, S. S. Sadat, and A. Shahbahrani, "Detecting, identifying, and counting vehicles based on deep learning algorithms in video surveillance systems," *Visual Journal of Technical and Vocational Education*, vol. 1, no. 2, pp. 79–90, 2024, doi: <https://doi.org/10.48301/jear.2024.459050.1027>.