Machine-Learning-Assisted Leak Detection Using Distributed Temperature and Acoustic Sensors

Gemeinhardt, Hayden; Sharma, Jyotsna

Published in:
IEEE Sensors Journal

DOI:
10.1109/JSEN.2023.3337284

Publication date:
2024

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Abstract—Small leaks, many of which often go undetected using conventional gauges, remain an urgent problem for an aging pipeline infrastructure. This article proposes a method that allows for an automated and robust leak detection and localization system by fusing distributed acoustic sensor (DAS) and distributed temperature sensor (DTS) data for machine learning. Distributed fiber-optic sensing creates an advantage over conventional gauges by providing real-time continuous measurements along the entire length of the fiber-instrumented pipeline at a high spatiotemporal resolution and sensitivity that can detect small leaks, as well as the leak location. High sensitivity, however, can create noisy data. Thus, machine learning is applied for a robust method of distinguishing nonleak data from leak signatures for accurate leak detection and localization. The workflow is demonstrated on an experimental pipeline setup exposed to environmental noise. The results illustrate reliable detection of small leaks between 0.04 and 0.30 L/s with F1 scores over 0.9, on a range of DAS frequency bands combined with DTS data. The fusion of two different types of distributed fiber-optic sensors not only increases the likelihood of small leaks being detected, but also decreases the number of false alarms by filtering noise that only appears in one sensor domain. Furthermore, the machine-learning approach utilizes segmentation, allowing for precise localization and quick investigation, as well as AI explainability of the sensor data deemed a leak signature.

Index Terms—Distributed fiber-optic sensing, image segmentation, leak detection, machine learning, sensor fusion.

I. INTRODUCTION

There are over two million miles of pipelines in the world, with over two-thirds contained in the U.S. alone [1]. Unfortunately, there have also been about 10,000 oil-and-gas pipeline failures to date in the U.S. since 2002 [2], and over half of the pipelines in the U.S. are over 40 years old [3]. Although pipelines are often seen as the safest and most economical method of transporting fluids, pipelines age over time, causing corrosion and degradation that opens small cracks and holes. Small leaks are more difficult to detect using traditional methods, such as pressure gauges and flow meters, due to their limited sensitivity and capability to only measure at discrete locations, which is often insufficient to monitor long pipelines. However, even small leaks can accrue over time resulting in severe environmental contamination and ecological and economic damage. For example, in 2006, the North Slope in Alaska had a spill of over 200,000 gallons of oil which went undetected for five days, originating from a hole only a quarter inch in diameter. It was only found when a worker noticed a petroleum odor when driving along a deserted road [4].

An aging pipeline infrastructure and numerous environmental disasters caused by pipeline leaks have spurred research in pipeline leak detection—both in enhancing existing sensor technology and in developing novel computational techniques for dissecting information from such sensors for creating reliable alerts.

A common approach to leak detection is to use pressure gauges or transducers at different locations throughout the pipeline [5]. A leak will cause a change in
pressure—depending on the leak rate—near the leak location, which then triggers an alert at the nearby pressure sensors based on the threshold set. Pressure sensors can also localize the leak using negative pressure waves [6]. However, the ability of pressure gauges to detect and localize leaks will depend on the location, sensitivity, and the number of gauges installed. The farther away the sensor is from the leak or the smaller the size of the leak, the less likely the sensor is to detect the leak.

Another proposal is to use acoustic point sensors [7], [8], [9], which can detect leak-induced vibrations. Acoustic sensors have the advantage of high sensitivity, allowing them to detect very small leaks (<0.001 L/s) [10]. However, acoustic point sensors have similar disadvantages as pressure sensors in terms of providing measurement at one location. Furthermore, the sensitivity of the acoustic sensors makes them especially prone to noise from the environment [8].

In response to the limitations of traditional point sensors and gauges, many studies have investigated the use of distributed fiber-optic sensors installed on pipelines for detecting leaks [10], [11], [12], [13], [14], [15]. In this technology, laser pulses are sent into an optical fiber, and the backscattered light signals are used to detect temperature changes (in the case of a distributed temperature sensor or DTS) and vibration changes (in the case of a distributed acoustic sensor or DAS), resulting from leaks [16]. These sensors have the benefit of providing spatially and temporally continuous information with a single lead cable without the need for any electronics at the sensed location. The fiber functions both as a sensor as well as a low-loss, high-speed conduit for data providing simultaneous measurement at high spatial resolution along the entire length of the pipeline. Their small size (<1 cm outer diameter), lightweight (therefore, fast response time), low data losses, and form factor versatility enable relatively easy installation on several kilometers of surface, subsurface, and subsea pipelines, with minimal disruption to flow. Fiber-optic sensors are immune to electromagnetic interference, operable at high temperatures and pressures, chemically inert, and corrosion-resistant. Additionally, since the sensing modality is based on optical modulation and the sensed information is carried using light, fiber-optic sensors enable both more precise localization and detection of signals, such as small leaks that often go undetected by conventional gauges [17], [18].

The implementation of sensors into pipelines for the purpose of leak detection is only as beneficial as their ability to successfully detect leaks while minimizing false alarms. For sensors to detect small leaks, they must be sensitive to small changes in the environment. However, high sensitivity leaves sensors vulnerable to noise that can resemble or become indistinguishable from the small leak signals. A simple threshold will not be robust enough to be useful in this situation due to the inherent variability in signal and noise intensities for different operational and environmental conditions. A threshold too high will miss small leaks that are often too difficult for traditional methods, whereas a threshold too low will be sensitive to noise and cause too many false alarms to be relied upon. Furthermore, the implementation must be realistic and cost-effective. This is, perhaps, the most limiting aspect that prevents pipes from being loaded with traditional sensors to counter the factor of distance between leak and sensor.

In recent years, many machine-learning approaches have been proposed to ameliorate these problems with leak detection [5], [6], [9], [10], [11], [13], [15], [19], [20]. The most challenging aspect is the proper balance between the type of sensor data and the type of model, especially for small leak detection. As described in Section I-A, the existing literature on the detection and localization of small leaks (<0.3 L/s) in a realistic pipeline scenario is limited. In this study, we propose the implementation and simultaneous data collection using DAS and DTS, and the use of machine learning for automatic leak detection and localization on an experimental pipeline setup. We investigate the accuracy of leak detection with only the DAS and only the DTS data and demonstrate that having simultaneous DAS and DTS data acquisition—which is easily achievable with commercially available optical cables with multiple fibers—significantly enhances the accuracy of detecting leaks as low as 0.04 L/s. This can be intuitively inferred as most noise that affects DAS will not affect DTS, and vice versa. Furthermore, our study demonstrates segmentation of predicted leaks—a step toward explainable artificial intelligence (AI) in pipeline leak detection such that the exact section of data within the data sample causing an alarm is highlighted for human analysis. The novelties of our approach, when compared to the existing literature on relatively small leak detection and localization, are described in the next section.

A. Literature Review and Novelties of This Study

The topic of pipeline leak detection has been widely explored in the literature, with many methods demonstrating efficient and effective leak detection and localization [21], [22], [23]. However, relatively small leaks still remain a problem, namely, leaks small enough to be undetectable or indistinguishable from noise via conventional sensors and gauges. Machine learning has gained traction in this field, due to its robustness against noise [24]. It has a unique advantage over traditional methods in being able to dissect and classify not only the difference between a leak and noise, but also between a leak and noise resembling a leak, thus allowing sensors to be more sensitive to small leaks that would otherwise be analogous to noise. While a major disadvantage of machine learning is the extra computational resources needed to monitor the pipeline compared to the traditional methods, the ability to detect smaller leaks could outweigh the cost.

In the realm of anomaly detection, machine-learning models have been demonstrated to quickly and intelligently process sensor data. CNNs particularly have been demonstrated in industrial settings for fault diagnosis of bearings [25], roadside obstacle detection in automated driving [26], [27], [28], structural health monitoring for large-scale civil structures [29], [30], [31], and abnormality detection in medical imaging of body organs [32], [33], [34]. The practicality of machine learning in sensor data is proven, yet the difficulty derives from applying a proper composition of sensor types, data
preprocessing and machine-learning models, pipeline setups, target leak rates, and so on, for each problem. In the applications concerning leakage detection, a vast array of possible solutions has been explored.

In the field of traditional sensors, many studies investigate the use of pressure and flow sensors. Zang et al. [5] used synthetic data generation to create large datasets based on pressure data and then combined four machine-learning models together via unified feature extraction to detect leaks as small as 0.388 L/s. However, they had no mention of detecting the leak location and did not investigate leaks smaller than 0.388 L/s. Mounce et al. [19] used an artificial neural network to compare predicted flow rates versus actual flow rates in a real-life water distribution pipeline using flow sensor data and provided an example where they detected a leak as low as 0.3 L/s. However, all of the other leaks were much larger, and given that there was also no analysis of leaks that were missed, the reliability of small leak detection and the ability to detect leak location is unknown. Spandonidis et al. [20] used long short-term memory autoencoders to monitor accelerometer data for spikes, then if a spike in measurements above the leakage threshold was detected, they generated a spectrogram and classified the image using CNNs. They achieved an accuracy of over 96% for a leakage diameter of 5 mm at a sensor distance of 33.5 m; however, the leak rate was not reported in the study.

Juliano et al. [7] used acoustic emission sensors to detect leaks 1.3–20.8 mL/s with a flow rate of 11.4 L/s, although it was on an underground pipeline away from environmental noise. Furthermore, many of the smaller leak rates were not detected, and the study had no mention of nonleak test cases. With acoustic data particularly, false alarms are a big problem when dealing with small leaks. It is easy to neglect false positives to achieve high results. Kouziopoulos et al. [8] used statistical analysis on acoustic data with a focus on detecting leak locations where the leak diameter ranged from 3 to 7 mm, but they also had no mention of nonleak cases in the results section. Quy and Kim [9] applied adaptive thresholding to acoustic data for leak localization, for a leak diameter of 0.3 mm and a pipe flow rate of 10 m³/h (2.7 L/s) on an indoor pipeline with sensors spaced only 2.6 m apart. Wang and Gao [6] proposed using a CNN model on negative pressure wave and acoustic wave fusion, but they did not specify leak rates and their experimental setup consisted of a 40-mm-diameter pipeline with a flow rate of 0.14 L/s which is not analogous to a real pipeline.

The papers using traditional sensors and gauges all suffered from limitations due to the discrete nature of the conventional sensors. Viewing the current literature, traditional sensors are not currently adequate enough to detect and localize small leaks in a noisy environment. Distributed sensors, on the other hand, have achieved more notable results, especially due to their ability to measure across the length of the installed fiber.

Wu et al. [11] use a DAS implementation for leak detection, utilizing a CNN on the time-space data, and reported an accuracy of about 86% in detecting small leaks which were less than 5% of the pipe flow rate. Two small leak cases were investigated corresponding to 0.084 and 0.027 L/s leak rates; however, less than 61% true positive rate was reported for 0.027 L/s leak. Additionally, the DAS data was preprocessed in 128 frequency bands covering 0–5 kHz, with a window size of 204.8 ms, which may be computationally heavy for large-scale pipeline implementation and long-term monitoring due to the amount of data needed for the analysis.

Guerriero et al. [12] fused DAS data with internal leak detection (comparing measured sensor data with predicted model conditions based on computer simulation) to detect a leak as small as 0.037 L/s with 99% accuracy, but it was demonstrated on buried cables using a pressure washer rather than a pipeline. Yan et al. [10] used pipeline leakage noise standard deviation to quantify microleaks of 0.00039 L/s using DAS data. No analysis of false alarm rates on nonleak data was presented.

Kim et al. [13] used a CNN on spectrograms of Raman-based DTS data, although they only investigated high leakage water temperature (35 °C to 90 °C) where the temperature difference between the environment and the fluid is large. Xu et al. [14] used a dynamic threshold identification method on a Raman-based DTS to detect temperature differences for leaks down to 0.016 L/s. In their study, the cable is positioned beneath the underground pipe, away from ambient temperature effects and flow noise.

Wang et al. [15] proposed fusing both DAS- and Brillouin-based DTS data together and extracting a set of features—such as the average of the data and range of the time domain—to create a random forest model for classification of normal data, leak data, and interference data (external noise). However, there is no mention of the leak rate investigated or the localization of the leaks nor evidence of the simulated pipe being exposed to a noisy environment.

Among the current literature, there appears to be a gap in reliable, small leak detection and localization on a realistic liquid pipeline with a noisy dataset. The full encapsulation of these features is necessary for a real field application, thus compromising on one aspect renders an inapplicable proposal. To this end, this study investigates the application of DAS and DTS for the automatic detection and localization of small leaks (0.04–0.30 L/s) in an outdoor experimental pipeline using machine learning using a small training dataset. The application of the two distributed sensing technologies is explored both individually, as well as with simultaneous DAS and DTS acquisition using a single lead cable with single-mode and multimode fibers that are easily available commercially. The goal was to quantify any improvement in leak detection reliability with a multisensor approach. Furthermore, we demonstrate precise segmentation of the leak data and its corresponding location along the pipeline. Localization is crucial when dealing with false alarms—which will be nearly impossible to completely eliminate in a noisy environment—as it allows for quick verification of a predicted leak. Segmentation also specifically provides the model with the unique ability to convey which piece of data it believes to be representative of a leak. Although segmentation is a significantly more difficult task than classification alone, it is beneficial for an operator to be able to visually investigate the suspected section of data. This provides insights into the
model’s behavior, which is especially important in small leak scenarios where noise and leak data are often indistinguishable from a model. Due to problematic black-box behavior with machine-learning models (such as, in our case, nuisance alarms), explainable AI has become a hot topic in the field of machine learning. We believe that our study provides an important contribution toward explainable AI in pipeline leak detection and allows us to take full advantage of precise distributed fiber-optic measurements and a robust model.

II. EXPERIMENTAL SETUP AND DATA COLLECTION

Experiments were conducted utilizing an outdoor experimental flow loop consisting of a 6.1-m-long stainless steel pipe of 10.16 cm outer diameter [35]. A tactical tight-buffered cable with single-mode and multimode fibers was helically wrapped around the pipe with an average pitch of about 1.88 cm and an effective length of 106.68 m, as shown in Fig. 1(a). The tactical cable has an outer diameter of 6.4 mm, with eight tightly buffered fibers within an aramid yarn reinforcement and a polyurethane outer jacket. The cable is rated for operating temperatures between -46 °C and 85 °C. Experiments were conducted by flowing water through the main pipe at different flow rates, with one or both leak ports open to create controlled leaks, as shown in Fig. 1(b). The main pipe is instrumented with a flowmeter to measure the flow rate that is controlled by a variable-frequency drive (VFD) connected to a centrifugal pump that is used for circulation. Leak rates were measured manually. The leaks occurring with the pump off were created by using the hydrostatic pressure differential between the water tank [used for circulating the water, as shown in Fig. 1(a)] and the pipe.

For DAS acquisition, a commercially available optical interrogator was used that uses Rayleigh-based backscatter to derive optical phase shift. The DAS data was analyzed using a frequency-domain signal-processing approach by implementing frequency band extraction (FBE) in multiple frequency ranges, as described in previous studies [18], [36]. The goal of the FBE approach is to investigate the signal content within the different frequency ranges with the objective of isolating the signals of interest from environmental noise in the frequency domain. DTS was acquired using a Raman backscatter-based commercial interrogator. The acquisition parameters for DAS and DTS measurements are summarized in Table I.

DAS and DTS data were collected on several different days during the different experimental trials. Since the focus of this study is the detection and localization of small leaks, the leak rates varied from 0.04 to 0.30 L/s, while the main pipe flow rate varied from 0 to 3.7 L/s. At 0 L/s main pipe flow rate, a varying leak condition was created using the hydrostatic pressure difference between the pipe and the water tank used for circulation. In this case, the leak continuously decreased until the tank water level equilibrated with the pipe height. The experimental dataset resulted in roughly 20 h of DAS and DTS data, for 58 leak signatures.

A. DAS Preprocessing

While the main advantage of a distributed sensor is the data acquisition at high spatial and temporal resolution, it results in voluminous streaming datasets that are multidimensional and complex [16]. Thus, the raw time-domain data is commonly transformed into the frequency domain using signal-processing techniques, such as the FBE method, which averages the data over a fixed time duration and frequency range for a simplified interpretation. The signals of interest can be more easily isolated from the background noise and environmental effects in the frequency domain [18]. Thus, the FBE approach can help with improved signal detection while also reducing the size of the DAS data being analyzed, resulting in computational efficiency.

As described in Sharma et al. [18], the FBE data is computed by taking the continuous stream of vibration data from the DAS interrogator, \( x(t, d) \), and splitting it into frames \( x_k(t, d) \) where \( t \) is the time, \( d \) is the depth, and \( k \) is the frame number. Then, a fast Fourier transform is applied to each frame of the vibration data \( x_k(t, d) \) to obtain the spectral information \( X_k(f, d) \). For each frequency band, \( b \), defined by its lower \( f_0(b) \) and upper \( f_1(b) \) frequency limits, the FBE profile is calculated as follows:

\[
FBE_k(b, d) = \frac{2}{N_{FFT}} \sum_{f=f_0(b)}^{f=f_1(b)} |X(f, d)|^2.
\]  

(1)

The number of bands and their respective range of frequencies vary by the specific application and the signals of interest. It is typically determined during an initial calibration of the

---

**TABLE I**

<table>
<thead>
<tr>
<th></th>
<th>DAS</th>
<th></th>
<th>DTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical mode</td>
<td>Single</td>
<td>Optical</td>
<td>Multimode</td>
<td></td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>1.5 m</td>
<td>Spatial</td>
<td>1 m</td>
<td></td>
</tr>
<tr>
<td>Spatial sampling</td>
<td>0.82 m</td>
<td>Sampling</td>
<td>0.5 m</td>
<td></td>
</tr>
<tr>
<td>Temporal sampling</td>
<td>10 sec</td>
<td>Sampling</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>16 km</td>
<td>Range</td>
<td>15 km</td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>10 kHz</td>
<td>Accuracy</td>
<td>± 0.5 °C</td>
<td></td>
</tr>
<tr>
<td>Gauge length</td>
<td>1.5 m</td>
<td>Operating Range</td>
<td>0 to 50 °C</td>
<td></td>
</tr>
</tbody>
</table>
fiber response in the specific operating environment (such as fiber installed on a pipeline, wellbore, storage infrastructure, etc.). In the absence of the initial calibration step, a general guideline is to divide the acquisition frequency range (up to the Nyquist frequency) into eight frequency bins, also known as an octave bin selection. The bins can then be further refined based on the fiber signal observations in the specific application. The reduction of bands is necessary for prompt analysis of the recorded data. For example, using a 1-s time frame and an FBE of ten bands, the data volume is reduced by 1000 [37]. The data reduction becomes especially important for real-time, continuous long-range sensing, such as that needed for monitoring pipelines that are several kilometers long. Further analysis can recognize and isolate bands that are either too noisy for practical use or can cleanly capture expected leaks in the matching frequency range. This, too, will depend on the application (such as the type of pipeline, flow rates, etc.).

The frequency bands utilized for DAS FBE processing in this study are summarized in Table II. Fig. 2(a) shows the collected leak samples from Band 0 of one of the datasets used for testing, which represents the full range of the acquired frequencies (0–2000 Hz), while Fig. 2(b) shows the FBE profile of the same dataset in all the different frequency bands to illustrate the effect of different frequency ranges on the signal and noise levels. DAS FBE in Fig. 2(a) indicates multiple leak signatures created at different main pipe flow rates at distinct time intervals at 61.74 m depth, which corresponds to the location of the leak port shown in Fig. 1(b). Bands 2–5 show the leak signal more clearly when compared to Bands 0 and 1, indicating a dominant frequency fingerprint of the acoustic or vibrational signal due to a leak above 50 Hz and less environmental effect (which includes ambient noise, such as the effect of wind and pump).

<table>
<thead>
<tr>
<th>FBE Band</th>
<th>Frequency Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 0</td>
<td>0-2000 Hz</td>
</tr>
<tr>
<td>Band 1</td>
<td>2-10 Hz</td>
</tr>
<tr>
<td>Band 2</td>
<td>50-200 Hz</td>
</tr>
<tr>
<td>Band 3</td>
<td>200-500 Hz</td>
</tr>
<tr>
<td>Band 4</td>
<td>1000-1500 Hz</td>
</tr>
<tr>
<td>Band 5</td>
<td>1500-2000 Hz</td>
</tr>
</tbody>
</table>

B. DTS Preprocessing

With the DTS temperature data, a gradient with respect to time was implemented to easily visualize the temperature changes resulting from the water leaks. The reasoning behind this is to account for environmental changes that are unavoidable, such as temperature differences between time of day, seasons, and weather conditions, without training a new model entirely for each situation. A Savitzky–Golay filter [38] was then applied to remove small noise spikes from the data, such that the leak data was better isolated from the environmental effects for the machine learning model. A Savitzky–Golay filter smoothens the signal by convoluting over the data and fitting adjacent points with a polynomial using a linear least-square method. Fig. 3 shows the resulting DTS gradient plots for a sample dataset that was used for the machine-learning model testing. The DTS in Fig. 3 and the DAS shown in Fig. 2 correspond to the same experimental trial used for machine-learning model testing. Different leak signals are observed in this dataset at the leak port location along the flow loop. The intensity of the observed temperature change varies due to the different pipe flow rates and leak rates tested. The leak rates and the corresponding main pipe flow rates are also shown in Fig. 3. It can be seen that the temperature difference resulting from a leak is less than 0.2 °C/s. This was intentional as it is relatively easier to detect a leak when there is high contrast between the ambient temperature and the temperature of the flowing fluid. In this study, we wanted to focus on leaks that are harder to detect using conventional approaches, such as those where the temperature change due to a leak is small.

III. MACHINE-LEARNING METHODOLOGY

DAS and DTS measurements are inherently different in terms of the data type, physical characteristics, acquisition rate, and processing, and thus each requires separate machine-learning approaches. To account for this, different models were used to segment each data type. One machine-learning model was trained using the DAS data, and another model was trained using the DTS data. For a prediction, a segment of the DAS and DTS data was sampled from the same time frame, fed into their respective models, and then the segmentation scores were added together and renormalized between 0 and 1 to be segmented at a threshold of 0.5, as outlined in the schematic in Fig. 4. The modeling approach used for each data type are described in detail in their respective sections below.
Fig. 3. DTS gradient data (in °C/s) with labeled leak rates (bottom) and main pipe flow rates (on top) of a sample dataset used in testing corresponding to the same DAS data shown in Fig. 2(a) and (b).

Fig. 4. Proposed fusion method in which DAS is segmented by U-Net, DTS is segmented by DFM, and the scores are added together for the final leak detection and localization.

In this study, the F1 score was primarily used to quantitatively measure the leak detection performance of the models. An F1 score is a useful metric for unbalanced datasets as it combines precision and recall into a single metric. Accuracy (or ACC) is also provided in the results as another reference. The pixel area under the curve (AUC) is used for preliminary testing in model selection for DTS data. The pixel AUC is calculated by mapping true-positive rates and false-positive rates for pixels at a range of thresholds between 0 and 1 and then taking the area under the resulting curve [39]. The F1 Score and accuracy can be calculated by (2c) and (2d), respectively, using the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values as

\[
\text{Precision} = \frac{TP}{FP + TP}, \quad (2a) \\
\text{Recall} = \frac{TP}{FN + TP}, \quad (2b) \\
\text{F1 Score} = \frac{2 \times (\text{precision} \times \text{recall})}{\text{precision} + \text{recall}}, \quad (2c) \\
\text{Accuracy (ACC)} = \frac{TP + TN \times 100}{TP + FP + TN + FN}. \quad (2d)
\]

The predicted leak location is intuitively determined to be the middle of the predicted segmentation. Since the cable was wrapped helically, the conversion from wrapped length to lateral length \(L\) was done using as follows, where \(H\) is the length of the cable, \(P\) is the pitch between wraps, and \(D\) is the diameter of the pipe:

\[
L = \frac{H}{P} \times \sqrt{(\pi \times D)^2 + P^2}. \quad (3)
\]

Then, the localization error was calculated via the absolute value of the difference between the leak port location and the mean value of the predicted segmentation along the depth axis.

A. DAS Data Analysis

The task of segmentation on DAS data is particularly difficult due to the high sensitivity of DAS that allows it to detect small leaks but makes it prone to noise. This is especially true in an exposed environment, as is the case in our experiment conducted in an outdoor flow loop in this study. As CNNs are some of the most successful and commonly used architectures in machine learning [40] and have demonstrated results in pipeline leak detection in previous studies [6], [11], [13], [20], the CNN-based model was utilized for the segmentation task. Specifically, a U-Net model was utilized. The U-Net architecture was first proposed by Ronneberger et al. [41] for biomedical image segmentation. U-Net is a widely popular architecture that gives the ability of segmentation with the advantages of a simple network structure, good performance, and a small training dataset size requirement [41].

U-Net is a network that consists of two parts: a contracting path with four encoder blocks and an expansive path with four decoder blocks, as demonstrated in Fig. 5. An encoder block consists of two layers of a 3 × 3 convolution followed by a rectified linear unit (ReLU) activation function and then a 2 × 2 max-pooling layer. A decoder block consists of a 2 × 2 transpose convolution layer then two layers of a 3 × 3 convolution followed by a ReLU activation function. Each of the encoder blocks is connected via a “bridge” and concatenated to the inputs of the respective decoder blocks. This causes the architectural representation to fold into a “U.”

U-Net is similar to autoencoders, which are designed to encode, or compress, an image into a feature vector and then decode, or reconstruct, the original image. The network is trained to learn patterns, such as leak data and normal data, and can then output those patterns to segmented masks. However, the skip connections in U-Net overcome a bottleneck in simple.
autoencoders where features can get lost in the downsampling and allow contextual information to get transmitted into the higher layers during upsampling.

A supervised learning approach was used with the model, such that leaks in the input data were labeled with a mask. The outputs of the model were confidence values between 0 and 1 of each pixel either belonging to the leak class or the normal data (or no leak) class. A threshold of 0.5 was then applied to convert the confidence values to binary values for segmentation and leak prediction.

Each input sample was set at 256 × 256 pixels, as the recorded length was easily adjusted to 256 pixels with the spatial and temporal resolutions of the acquired fiber data. Hyperparameter tuning found an optimal learning rate of 0.0001 and batch size of 32. The number of layers and filters are also tunable with U-Net architecture. This study used 32 filters over four layers since further increasing the complexity resulted in a longer computation time with minimal benefit.

### B. DTS Data Analysis

The leak initiated at the leak port results in wetting of the adjacent fiber wrapped near the leak port which creates a broad temperature change signal in the DTS gradient. Thus, unlike DAS, it was more difficult to create leak masks on the resulting data. Consequently, DTS modeling required a more dynamic and flexible approach, allowing the model to learn to generalize and detect anomalous areas itself as opposed to learning strict boundaries set in discrete locations.

Unsupervised anomaly detection has become a popular topic in the field of computer vision to address similar issues. Anomaly detection algorithms are typically designed to learn patterns or generalizations of normal data such that when anomalous data is introduced, and the model can detect the broken pattern and label the anomaly. To implement anomaly detection architectures, the Anomalib Python package was used. Anomalib was introduced by Akcay et al. [39] as an open-source library that comprises state-of-the-art anomaly detection algorithms. The library’s modularity makes it easy to test new algorithms and collect quantifiable results in a uniform manner.

To select the best model for further in-depth investigation among the different anomaly detection models available in the Anomalib package, a small subset of DTS data was trained and tested using a variety of different algorithms, and the results are summarized in Table III.

![Fig. 6. DFM architecture flow used for DTS leak prediction.](image)

The results show that many of the algorithms failed to generalize over the DTS data, such as Draem, CFA, and CFlow, based on the high FP or FN rates. Furthermore, while all of the remaining algorithms achieved perfect classification results, the pixel AUC varies between models. Pixel AUC is calculated from the predicted mask and a ground-truth mask that was put over the anomalous DTS data. The ground-truth mask is not used in training and is merely used to test the general capabilities of the models. Draem, Reverse Distillation, and CSFlow all had subpar AUC scores. Deep feature modeling (DFM) achieved the highest AUC score and overall performance.

<table>
<thead>
<tr>
<th>Model</th>
<th>Pixel AUC</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coupled hypersphere-based Feature Adaptation (CFA) [42]</td>
<td>0.82</td>
<td>Leak 25/25, No-Leak 0/25</td>
</tr>
<tr>
<td>Conditional Normalizing Flow (CFlow) [43]</td>
<td>0.85</td>
<td>0/25, 25/25</td>
</tr>
<tr>
<td>Cross Scale Normalizing Flow (CSFlow) [44]</td>
<td>0.58</td>
<td>25/25, 24/25</td>
</tr>
<tr>
<td>Deep Feature Modeling (DFM) [45]</td>
<td>0.88</td>
<td>25/25, 25/25</td>
</tr>
<tr>
<td>Discriminatively-trained Reconstruction Anomaly Embedding Model (DRAEM) [46]</td>
<td>0.69</td>
<td>25/25, 0/25</td>
</tr>
<tr>
<td>Patch Detection Modeling (PDNM) [47]</td>
<td>0.83</td>
<td>25/25, 25/25</td>
</tr>
<tr>
<td>Reverse Distillation [48]</td>
<td>0.69</td>
<td>25/25, 25/25</td>
</tr>
<tr>
<td>Student-Teacher Feature Pyramid Matching (STFPM) [49]</td>
<td>0.80</td>
<td>25/25, 25/25</td>
</tr>
</tbody>
</table>

Therefore, it was selected for further implementation and testing with the full dataset.

The DFM model implementation used in this study (see Fig. 6) consists of a feature extraction stage that uses a ResNet50 [50] backbone pretrained on the ImageNet dataset [51]. The feature extraction outputs a semantic feature vector that is then fed into a principal component analysis (PCA) model. Finally, a threshold was applied to the feature reconstruction error score for segmentation.

### IV. RESULTS AND DISCUSSION

This section compares the leak detection and localization results using machine-learning models with 1) DAS individually; 2) DTS individually; and 3) DAS and DTS data combined.

The respective machine-learning models were trained on randomly selected samples from 14.5 h of training data containing both leak and nonleak signals. The blind testing was performed using randomly selected samples from two datasets of simultaneous DAS and DTS data which were not used in training, containing 15 leaks each lasting 10–20 min each over the course of 5.5 h, along with 2 h of no leak data at various main pipe flow rates. The testing used a total of 100 randomly selected samples from the data of each band of DAS and the data of DTS. Of these, 82 samples contained leaks and 18 samples contained no leak signals.

#### A. Performance Using Only DAS

The model performance results for both leak detection and localization using only the DAS data corresponding to
the different FBE bands are summarized in Table IV. Band 0 encompasses all the acquired frequencies [0–2000 Hz], while the other bands represent isolated frequency ranges that demonstrate different levels of signal and noise effects [as illustrated earlier in Fig. 2(b)]. Each band was tested separately and the resulting accuracy and F1 scores for each band are presented, along with the localization error. For in-depth analysis, the true and false positive and negative counts (TP, FP, TN, and FN, respectively) are also shown.

It is observed that Bands 0 and 1 performed worse than the other bands, which was expected as they were plagued with noise that masked the smaller leaks (<0.1 L/s), which was also seen earlier in Fig. 2. However, the larger leaks (>0.1 L/s) were still distinguishable in the noisy bands. Band 5 had a surprisingly high FP rate compared to Bands 3 and 4. Upon analysis, this band was most prone to an unexpected “column of noise,” which appeared each time there was a change in the flow rate in the pipe, likely contributed by a pump vibration as the flow rate changed. These columns can be seen in examples in Fig. 2(b). However, these FPs were only small spots of segmentation that could be relieved by requiring an alarm to be clusters of, for example, 20 pixels or more, depending on various cost functions.

Due to the vast differences in accuracies between bands, frequency analysis should be performed to determine the best bands before leak monitoring in a real application during an initial fiber response calibration process. This can be done using spectrum and spectrogram analysis, as demonstrated in previous studies by the authors [52], [53]. The frequency ranges corresponding to the leak and the noise may not be the same between pipelines and fiber deployments, as each pipeline will have different flow rates, types of transported fluids, the diameter of the pipe, thickness of the pipe, type of pipe material, sources of noise, fiber installation method, fiber material, and so on.

Across all bands, there was an average localization error of about 2.98 m in terms of the length measured along the fiber-optic cable. However, because the fiber is wrapped helically on the pipe, the localization error is actually lower in terms of the true lateral distance which was estimated to be about 0.17 m using (3). The results in Table IV report the localization error in terms of the converted lateral distance. Additionally, the leaks with the highest error were the larger leaks, as the acoustic signal was skewed toward the location where the leaking water landed on the pipe which was a littleoffset to the leak port location. This phenomenon is illustrated with an example in Fig. 7. Fig. 8 presents three examples of leak detection and localization using only DAS data to illustrate the performance of the model. In each figure, the first picture shows the DAS input segment, including its corresponding time and depth locations, the second picture shows the true location of the leak as indicated by the yellow-colored region, the third picture shows the DAS prediction scores (on a scale of 0 to 1, where 1 indicates high probability of leak), and the fourth picture shows the predicted leak based on a threshold selection of 0.5.

In Fig. 8(a), there are two leaks corresponding to 0.10 L/s leak rates. The DAS model successfully detected both leaks with localization errors of 0.06 and 0.12 m. (b) DAS sample with an FN and an FP, missing the beginning of a 0.09 L/s leak and falsely predicting noise as a leak. (c) DAS sample with an FN, missing the beginning of a 0.10 L/s leak. The TP leak was 0.04 L/s with a localization error of 0.07 m.

### Table IV

<table>
<thead>
<tr>
<th>Band</th>
<th>ACC</th>
<th>F1 Score</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Avg. Localization Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>83%</td>
<td>0.89</td>
<td>68</td>
<td>3</td>
<td>15</td>
<td>14</td>
<td>0.17 m</td>
</tr>
<tr>
<td>1</td>
<td>83%</td>
<td>0.90</td>
<td>75</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>0.14 m</td>
</tr>
<tr>
<td>2</td>
<td>87%</td>
<td>0.92</td>
<td>78</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>0.16 m</td>
</tr>
<tr>
<td>3</td>
<td>97%</td>
<td>0.98</td>
<td>81</td>
<td>2</td>
<td>16</td>
<td>1</td>
<td>0.17 m</td>
</tr>
<tr>
<td>4</td>
<td>98%</td>
<td>0.99</td>
<td>80</td>
<td>0</td>
<td>18</td>
<td>2</td>
<td>0.19 m</td>
</tr>
<tr>
<td>5</td>
<td>87%</td>
<td>0.93</td>
<td>82</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>0.18 m</td>
</tr>
</tbody>
</table>
seen as the acoustic noise increases (indicated by the transition from a darker to a brighter color) along the Y-axis. Fig. 8(b) shows an example where the DAS model did not successfully detect the start of a leak of 0.09 L/s rate. A likely reason for the missed leak is the model predicted the leak to be noise. A possible solution is to increase the DAS temporal resolution to collect more data so that the model can differentiate between noise and leak data effectively. However, this problem is alleviated by fusion with DTS, as discussed in the later section.

Fig. 8(b) also has a small segment predicted to be a leak which was not a true leak. The two errors in the example demonstrate two flaws in the model using only DAS data: 1) there need to be more training samples when the leak signature present in the segment is for a very short duration and 2) there need to be training samples to account for the noise signatures, such as the “columns of noise” throughout the examples resulting in higher than normal prediction scores for nonleak data. In this analysis, one can easily see the benefits of a model selection that allows for AI explainability, so that one can learn its behaviors and flaws to reduce future false or missed alarms.

Fig. 8(c) illustrates an example of detecting a leak as low as 0.04 L/s with a localization error of 0.07 m. In the same figure, the beginning of a small leak (at 12:02:05) was missed. This is similar to the discussion of Fig. 8(b).

**B. Performance Using Only DTS**

The model performance results using only the DTS data are summarized in Table V. The anomaly detection algorithm worked exceptionally well at segmenting leak data; however, it had a very high FP rate. This is because the DFM model does not distinguish between noise and leak signatures, but simply classifies anomalies. The approach proves very beneficial for quick detection of leaks, although the FPs make the model unreliable on its own. However, the choice of the approach was not made with the intention of its use individually, but in combination with DAS data, which was hypothesized to solve the FP issue with DTS data and take advantage of the approach under the intuition that it is unlikely temperature noise will occur at the same spot as acoustic noise and be wrongly predicted by both models.

The localization error was larger on DTS data than on DAS data. Although, intuitively, the center of the segmentations would still overlap, the nature of the DFM model results in less strict segmentation boundaries, thus skewing it a few pixels in one direction or the other.

It is worth noting that in this experiment, there was a small temperature difference between the environment and the liquid inside the pipeline (<0.2 °C/s), as discussed previously. A pipeline that transfers liquid with much higher or lower temperatures than the environment would demonstrate more benefit (and potentially better results) with the DTS data.

**C. Performance After DAS and DTS Fusion**

The model performance results for DAS and DTS data fusion are summarized in Table VI. As hypothesized, combining both DAS and DTS data gave better results than DAS or DTS individually. For DTS, the high FP rate significantly improved. For DAS, the results varied by frequency band.

**TABLE V**

<table>
<thead>
<tr>
<th>ACC</th>
<th>F1 Score</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Avg. Localization Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>87%</td>
<td>0.93</td>
<td>83</td>
<td>13</td>
<td>4</td>
<td>0</td>
<td>0.46 m</td>
</tr>
</tbody>
</table>
Band 0’s FN and Band 5’s FP dropped over 50%, and Band 4’s accuracy increased to a perfect 100%. Bands 3 and 4 still outperformed the other bands in fusion, which is similar to the observation with model predictions using DAS alone (see Table IV) for those frequency bands. This is expected as the leak signal had the lowest noise and highest intensity in that frequency band, which is specific to the data being analyzed in this case. In other cases, the frequency with the highest SNR of the signal of interest and lowest noise may differ. This supports the idea that preprocessing the DAS data in different frequencies can further improve the machine-learning model results. Additionally, analyzing these irregularities can help to tailor the model accordingly. Overall, the small leak detection performance noticeably improved when fusing both DAS and DTS predictions together.

The localization accuracy was similar to that of the DAS individually. Some localization errors were higher since the DTS predictions had much higher errors than DAS, therefore resulting in a slightly increased error when the two are combined. However, the increases compared to DAS alone are mostly negligible for practical purposes, with a maximum increase of 0.01 m. The average DTS localization error, on the other hand, decreased by 60%–70% when fused, from 0.46 m to as low as 0.15 m.

Fig. 10 shows examples of the predictions that highlight and discuss various advantages of fusion. The picture in the first column shows the DAS sample collected from the data. The second column shows the DTS sample collected from the same time frame. The third column shows the ground-truth leak location (represented by the yellow-colored region), while the fourth and fifth columns show the results after combining the DAS and DTS predictions and then applying a threshold of 0.5 to segment the prediction, respectively.

Table VI contains the band-by-band performance of the DAS + DTS model prediction fusion.

<table>
<thead>
<tr>
<th>DAS Band + DTS</th>
<th>ACC</th>
<th>F1 Score</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Average Localization Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 0+DTS</td>
<td>91%</td>
<td>0.91</td>
<td>76</td>
<td>3</td>
<td>15</td>
<td>6</td>
<td>0.18 m</td>
</tr>
<tr>
<td>Band 1+DTS</td>
<td>86%</td>
<td>0.91</td>
<td>74</td>
<td>6</td>
<td>12</td>
<td>8</td>
<td>0.15 m</td>
</tr>
<tr>
<td>Band 2+DTS</td>
<td>92%</td>
<td>0.95</td>
<td>80</td>
<td>6</td>
<td>12</td>
<td>2</td>
<td>0.17 m</td>
</tr>
<tr>
<td>Band 3+DTS</td>
<td>98%</td>
<td>0.99</td>
<td>81</td>
<td>1</td>
<td>17</td>
<td>1</td>
<td>0.18 m</td>
</tr>
<tr>
<td>Band 4+DTS</td>
<td>100%</td>
<td>1.00</td>
<td>82</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0.19 m</td>
</tr>
<tr>
<td>Band 5+DTS</td>
<td>94%</td>
<td>0.96</td>
<td>82</td>
<td>6</td>
<td>12</td>
<td>0</td>
<td>0.17 m</td>
</tr>
</tbody>
</table>

Fig. 10(a) shows the fusion of the DAS and DTS samples in Figs. 8(a) and 9(a), detecting 0.10 L/s leak with localization errors of 0.06 and 0.13 m on top and bottom, respectively. (b) Fusing of the simultaneous DAS and DTS samples in Figs. 8(b) and 9(b), where DAS missed the leak and DTS had a false alarm. By fusing the data, DAS removes the DTS FP and DTS removes the DAS FN, with an improved localization error of 0.04 m. (c) Fusing of simultaneous samples as seen in Figs. 8(c) and 9(c), where DAS missed the leak and DTS had poor localization. By fusing the data, DAS corrects the localization errors to 0.07 and 0.09 m, respectively, and DTS corrects the DAS FN.

Fig. 10(b) shows the fusion of simultaneous DAS [see Fig. 8(b)] and DTS [see Fig. 9(b)] data in which the DTS model detected the 0.04 L/s leak and falsely detected a second leak due to temperature noise (in the middle), while the DAS model missed the 0.04 L/s leak and also falsely predicted a second leak due to acoustic noise in a different location than the DTS model. By fusing the two predictions together, the FPs and FN between the two predictions were filtered out. The final localization error was 0.04 m, drastically reducing the DTS localization error of 0.21 m as seen in Fig. 9(b). This example illustrates how noise can appear and be falsely predicted as leaks in two different domains (acoustic and temperature) and be successfully filtered out when combining the data. By doing so, a false alarm is only to occur when the noise occurs at the same time and location in both domains and both models also falsely predict the noise as a leak.

Fig. 10(c) illustrates another example of the fusion of DAS [see Fig. 8(c)] and DTS [see Fig. 9(c)] data in which a leak of size 0.04 L/s was detected by both models individually—although the DTS model had poor segmentation—and the beginning of a second leak of 0.10 L/s was detected by the DTS model but undetected by the DAS model. By fusing the two predictions together, the DAS FN was corrected and the DTS localization error was lowered from 0.08 and 0.12 m to 0.07 and 0.00 m, respectively. The example demonstrates how small leak detection and localization can be improved by combining two models such that their advantages complement each other and their disadvantages cancel out.
The results speak to the intuitive reasoning that combining sensors filters noise by requiring predicted leak signatures to be present in both domains. Furthermore, it shows that models for each type of data can be strategically chosen to leverage this fact. It is better that a model predicts noise as a leak than predicts a leak as noise since the noise predicted as a leak will likely be filtered upon fusion. Then, by leveraging the fusion advantage, the models can tackle the more difficult problem of segmentation for localization and explainability and do so specifically and accurately for small leaks. However, the downside of such an implementation is an increase in complexity in the computational processing and workflow and an increase in sensor data archival costs, in addition to the cost of the two interrogators needed for DAS and DTS acquisition. Thus, the technical, environmental, and economic considerations have to be weighed for each individual application.

The sources of noise apparent in the samples were only realized upon postprocessing. However, such unpredicted instances of noise are likely to appear in a real field scenario and thus only bolster our hypothesis. It is impossible to train a model to account for all types of noise from an environment and can often be indistinguishable from a leak.

While this study used a threshold of 0.5, which was optimally aligned with the F1 score, a real application would investigate different optimal thresholds based on real-life cost functions, ideally balancing both economic and environmental considerations.

V. CONCLUSION AND FUTURE WORK

This study investigated the application of machine learning for automatic leak detection with distributed fiber-optic sensing. Specifically, this study proposed the fusion of DAS and DTS data for small leak detection and localization via segmentation. By combining DAS and DTS data, this study demonstrated an improvement in the detection of specifically small leaks (<0.3 L/s), detecting a leak as low as 0.04 L/s. Furthermore, it illustrated the intuitive reasoning that fusion data reduces false alarms by filtering noise that only appears in one sensor domain and increases true alarms in cases where the leak is small enough to be missed by both sensor models individually but caught by combining results. Lastly, the problem of black-box predictions is addressed by utilizing segmentation, a step toward explainable AI in leak prediction.

There remain many possible avenues for improvement in the current study. While one of the focuses of this article was the detection of the leaks, the workflow can be extended for quantifying the leak rates by leveraging previous work by the authors [35], [36]. Among other things, there also remains the investigation into different fiber configurations other than helical wrapping, such as lateral placement of fiber outside the pipe, placement of the fiber inside, embedded in the pipe walls, or away from the pipe. Furthermore, in the future, the study can be extended to sensor data acquired from a longer pipeline, tested on a larger set of sensor data for different types of fluids (including gases and gas–liquid mixtures), and tested on additional leak rates.

ACKNOWLEDGMENT

The authors would also like to extend their gratitude to Ramji Tangudu and Jagadeeshwar Tabjula for their assistance with the experiments. We also acknowledge collaboration from Prof. Manas Gartia and Prof. Jianwei Wang.

REFERENCES
