NDT&E International 144 (2024) 103073

Contents lists available at ScienceDirect

NDT and E International

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

RHiREM: Intelligent diagnostic framework for pipeline Eddy Current Internal Inspection based on reinforcement learning with hierarchical reward exploration mechanism

Lanqin Su [a], Bin Gao [a,*], Xiangyu Zhao [a], Yang Fu [a], Wai Lok Woo [b]

[a] School of Automation Engineering, University of Electronic Science and Technology of China, China
[b] Department of Computer and Information Sciences, Northumbria University, England, UK

ARTICLE INFO

Keywords:
Pipeline internal detection
Automated detection
Eddy current time-series data
Deep reinforcement learning

ABSTRACT

Pipeline safety is of paramount importance for socio-economic development, necessitating regular inspection and maintenance. The complexity of the internal pipeline environment and the presence of irregular noise in detection data, however, pose significant challenges to pipeline inspection technologies. Additionally, the reliance on manual expertise for inspection and the absence of standardized assessment criteria further complicate the development of automated inspection methods in this context. To address these challenges, this paper proposes a pipeline anomaly detection framework based on reinforcement learning with hierarchical reward exploration mechanism (RHiREM). It includes two main aspects: Firstly, the hierarchical reward mechanism. By deeply simulating the process of defect recognition based on expert personal experience, the original pipeline data is first adaptively divided into sets of windows with different sizes, and then attribute and type judgments are performed on them. In this way, the approach achieves accurate identification of defects and pipeline structures in scenarios with minimal noise interference. Secondly, the hierarchical exploration mechanism. By leveraging the temporal exploration and spatial exploration, the mechanism enables further deep search and feature learning on complex pipeline signals, and facilitates comprehensive assessment of the relationships between global features and local features across different signals, effectively resolving the difficulties associated with identifying defect signals in the presence of high noise interference. the proposed framework has been demonstrated to automate the detection of complex on-site pipeline internal signals and successfully detected the common anomalies with high F1-score over conventional techniques.

1. Introduction

Pipelines serve as critical components of the infrastructure network, playing a crucial role in ensuring the safety, reliability, and efficiency of various industries, including oil and gas transportation. Regular inspection and maintenance are therefore essential and important. As a non-destructive inspection method, eddy current testing technology offers advantages such as high speed, low cost, and ease of automation, making it an excellent approach for rapidly detecting internal defects in pipelines. This method utilizes electromagnetic principles to detect and analyze defects in the pipeline structure, exhibiting a high sensitivity to small cracks, corrosion, and other internal anomalies. By analyzing the time-domain signals generated from the interaction between induced eddy currents in the pipeline and the existing defects, these anomalies can be effectively detected. The problem of pipeline defect identification based on eddy current signals essentially falls within the realm of time-series data anomaly detection, which has been extensively studied by numerous researchers across the globe.

In the research on single-method anomaly detection, various algorithms have been proposed, including K-means clustering, Isolation Forest (iForest), deep learning autoencoders, and time series clustering labeling. Gupta et al. [1] introduced a K-means clustering anomaly detection algorithm based on local search. It detects anomalies by clustering the dataset and analyzing the distribution differences of abnormal and normal samples in different clusters. Liao et al. [2] proposed the iForest anomaly detection algorithm, which uses dimension entropy as the partition criterion to construct the forest and compares the path length of the test samples to the average path length of normal

* Corresponding author.
E-mail address: bin_gao@uestc.edu.cn (B. Gao).

https://doi.org/10.1016/j.ndteint.2024.103073
Received 2 December 2023; Received in revised form 30 January 2024; Accepted 16 February 2024
Available online 21 February 2024

samples for anomaly detection. Wang et al. [3] presented a deep learning autoencoder anomaly detection algorithm that learns the low-dimensional representation of data using autoencoders and measures the anomaly level of samples based on reconstruction errors. Bu et al. [4] developed a time series clustering labeling algorithm that clusters historical time series and labels the actual centroids of each cluster. New time series are subsequently input into the labeled clustering model for real-time clustering and anomaly detection. Kha et al. [5] proposed a clustering and anomaly detection approach based on time segment partitioning. It divides the target time series into equally-sized subsequences, performs clustering on the subsequences, and determines whether a subsequence is an anomaly based on the clustering results.

In the research on multi-source methods for anomaly detection, the fusion of different algorithms significantly improves the accuracy and reliability of anomaly detection. Ren et al. [6] combined spectral residuals with Convolutional Neural Networks (CNN) for anomaly detection in medical images. Xu et al. [7] proposed a method that combines autoencoder reconstruction errors with bias for anomaly detection in web data streams. Zhao et al. [8] addressed the holiday effect in time series data by incorporating the modeled holiday impact into anomaly detection. Furthermore, there are fusion methods combining different algorithms, such as fusion of iForest and Local Outlier Factor (LOF) algorithms [9].

In the framework research of anomaly detection, Yu et al. [10] proposed a method that establishes a historical data model and uses prediction confidence intervals to detect whether target samples are anomalies. Malhotra et al. [11] utilized neural network models to learn pattern features of time series data and used prediction errors to identify anomalies, making it suitable for time series anomaly detection. Ferdousi et al. [12] introduced the Peer Group Analysis theory, which utilizes clustering methods to detect anomalous points and is applicable to high-dimensional data. Liu et al. [13] proposed a generative approach for active learning, significantly reducing the cost of labeled data. Takeuchi et al. [14] introduced a novel scoring strategy to assess the degree of deviation, which is applicable to various domains. Teng et al. [15] proposed a unified framework for anomaly detection that supports various anomaly detection algorithms and models. Liu et al. [16] presented an integrated method for anomaly detection in large-scale system logs, enabling fast and efficient anomaly identification.

Traditional statistical models are suitable for online processing but lack accuracy, while supervised models have higher accuracy but suffer from a lack of labels. In the fields of anomaly detection and time series prediction, the absence of labels is a common characteristic of datasets. To address this, researchers have employed various methods, including models such as Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost), as well as variations of LSTM. For time series prediction, Li et al. [17] proposed an integrated LSTM and XGBoost model for forecasting, which improves the prediction performance compared to a single model. In a comparative experiment for time series prediction, Yamak et al. [18] found that ARIMA performs better for small datasets, while LSTM and GRU excel at capturing data features and patterns in large-scale datasets. Additionally, researchers have proposed other methods, such as using the Symbolic Aggregate approXimation (SAX) method for subsequence synthesis and employing LSTM for prediction [19]. Regarding the comparison of model variations, Siami-Namini et al. [20] found that BiLSTM achieves better prediction results in time series prediction experiments. However, due to its bidirectional path computation, BiLSTM has slower convergence and longer computation time than LSTM. Furthermore, Greff et al. [21] tested eight variations of LSTM and found that none of them surpassed the standard LSTM architecture, indicating that the forget gate and output gate are the most critical components of LSTM. As different types of models have different advantages in anomaly detection and time series prediction, it is important to choose the appropriate model and method based on different datasets and task requirements.

The aforementioned methods, however, have limitations in the analysis of eddy current signals in field pipelines, which mainly result in the following two challenges: (1) High data noise: (a) Eddy current sensors are susceptible to magnetic field disturbances from various external equipment during signal reception, resulting in the introduction of noise; (b) The process of converting the magnetic signal received by the eddy current sensor into an electrical signal also introduces noise; (c) Other interferences, such as inherent noise from the eddy current sensor system and power frequency noise, can also contribute to the overall noise. Therefore, the acquisition of raw data from field pipelines using eddy current sensors introduces a significant amount of irregular random noise, which obscures the effective signals. Extracting valid defect signals from the noise is one of the current challenges in eddy current signal anomaly analysis. (2) Complex signal conditions: (a) Eddy current sensors are highly sensitive and can be influenced by factors such as detection speed, lift-off distance, detection angle, and defect size, resulting in complex-shaped defect detection data; (b) Original structural components of the pipelines (such as welds, bends, flanges, etc.) are clearly present in the defect detection data; (c) The operating speed and vibration level of the eddy current sensor in field pipelines significantly affect the defect detection data. Therefore, the eddy current detection data in field pipelines exhibit complex and diverse characteristics, making the defect detection process heavily reliant on manual experience and lacking a fixed standard for reference. Consequently, designing algorithms for automated defect detection becomes the second challenge in eddy current signal anomaly analysis.

To address the aforementioned challenges, this paper presents a pipeline internal inspection intelligent diagnostic framework based on a hierarchical reinforcement mechanism. The distinctive feature of this framework lies in its utilization of the hierarchical reward mechanism and the hierarchical exploration mechanism in reinforcement learning to simulate the process of human expert judgment of pipeline state signals. It employs a multi-dimensional perspective to assess signals with high noise interference or those difficult to identify artificially, thereby obtaining accurate identification results and achieving automation to some extent. The main contributions of this paper can be summarized as follows:

(i). Hierarchical reward mechanism: The entire pipeline internal inspection intelligent diagnostic framework is constructed based on a hierarchical reward mechanism in reinforcement learning. By simulating the defect identification process through deep expert-based experience, accurate identification of defects and pipeline components is achieved under low noise interference. Ultimately, this framework enables precise automated detection of internal defects in pull-in pipelines.

(ii). Hierarchical Exploration Mechanism: The entire intelligent diagnostic framework based on hierarchical reward mechanism is fine-tuned and optimized by the hierarchical exploration mechanism. By conducting deep searches and feature learning in the spatiotemporal dimensions of real-world pipelines' complex state signals, it enables a thorough evaluation of the global and local feature relationships among different signals. This approach resolves challenges such as the difficulty of identifying defect signals in the presence of significant noise interference. It enhances the efficiency of defect and pipeline component recognition in complex environments, ultimately achieving a more accurate automated detection of internal pipeline defects.

We term the proposed framework as Reinforcement Learning with Hierarchical Reward Exploration Mechanism (RHiREM). The remaining sections of the paper are structured as follows: Section II provides an overview of the eddy current detection system, including sensing system structure, eddy current data characteristics, and pipeline environment difference. Section III presents the overall construction and fine-tuning optimization of the pipeline internal inspection intelligent diagnostic

framework in the study. Section IV presents a comprehensive description of the experimental conditions, results analysis and significant findings. Finally, Section V summarizes the contributions made by this paper and provides an overview of potential directions for future research.

2. Related work

2.1. Sensing system structure

The pipeline internal inspection system [22] is a collection of sensors, including eddy current probes, inertial measurement units (IMU), and odometers, which are primarily used for collecting qualified data of the pipeline inner wall. The specific structure is shown in Fig. 1, where the eddy current probes are arranged in a circular distribution with a certain angular spacing. Typically, there are 20 probes, and each probe corresponds to a set of time-domain data, which includes two dimensions of amplitude and phase. Besides, the hardware structures are mainly composed of FPGA, MCU, AD/DA conversion, power amplifier, and amplitude extraction. The FPGA generates two signals through the DAC, one as the excitation signal, and the other as the reference signal. It extracts the amplitude and phase of the signal, and the ADC collects the signal after the extraction. The MCU stores the data sent by the FPGA and communicates with the host computer to complete the data storage and real-time display. Therefore, the format of the collected eddy current data by the pipeline internal inspection system is a multi-channel time-domain signal. In this paper, only the amplitude of eddy current data is used to identify the anomaly, other sensor data is employed to assist anomaly localization and pipeline status monitoring.

2.2. Eddy current data characteristics

Due to the interference from pipeline structural components (such as welds, flanges, elbows) or the vibrations of the inspection system caused by factors like blockages or uneven operating speeds, the pipeline internal inspection system has a significant impact on the corresponding defect signals during the data acquisition process, as depicted in Fig. 2 and Fig. 3.

Fig. 2 illustrates the interference of pipeline structural components on defect signals. The red window represents the signal from a structural component (e.g., weld), while the blue window represents the defect signal. It is evident that they are very similar in terms of shape and amplitude, making it challenging to effectively differentiate between them. Fig. 3 demonstrates the influence of self-induced vibrations of the inspection system on defect signals. The significant noise generated by vibrations can mask the smaller amplitude defect signals, rendering the identification of minor defects difficult. Therefore, these two factors are also the challenges faced in the anomaly detection of eddy current signals in this paper.

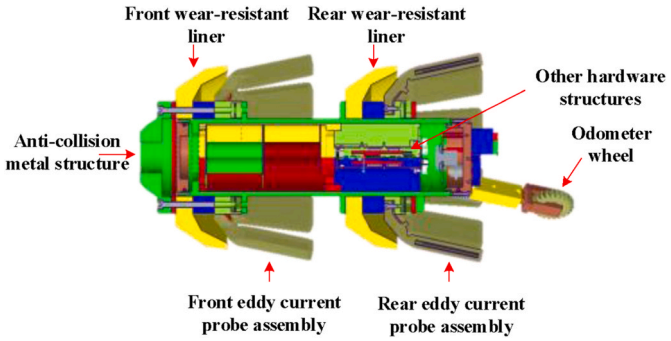


Fig. 1. Structure of pipeline internal inspection system.

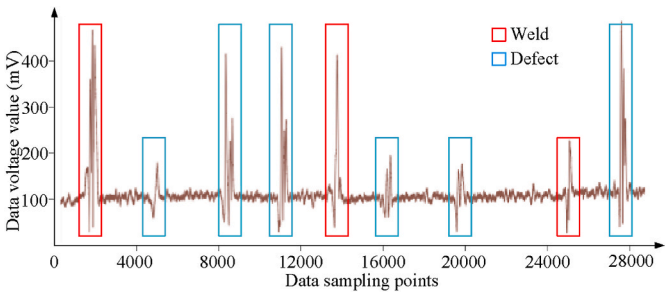


Fig. 2. Interference of pipeline structural components on defect signals.

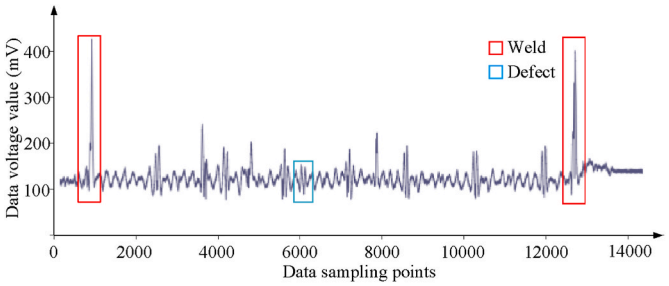


Fig. 3. Impact of pipeline inspection system vibrations on defect signals.

2.3. Pipeline environment difference

In addition, the specific application environments of the pipeline internal inspection system mainly fall into two categories: pull-in pipelines and field pipelines. As shown in Fig. 4, pull-in pipelines (Fig. 4(a)) refer to pipelines that have not been put into actual engineering use but contain artificially curated defects. They are typically used for preliminary algorithm validation or feature learning. Field pipelines (Fig. 4 (b)), on the other hand, are pipelines that are in actual engineering use and contain naturally occurring defects. They are primarily used to verify or enhance the algorithm's generalization and robustness. The specific differences between the two are illustrated in Table 1 (shown in appendix). The frame work is shown in Fig. 5.

3. Methodology

To date, non-destructive testing (NDT) techniques for oil and gas pipelines have made significant progress in the analysis and processing of eddy current signals. However, due to poor inspection environments, high signal noise, various pipeline conditions, complex defect characteristics, and insufficient accuracy of the detection systems, the quality of defect data deteriorates making subsequent signal analysis challenging and unable to meet the actual inspection requirements of all pipelines. Therefore, existing methods for analyzing eddy current signals in pipelines still have significant limitations. To address these challenges, this paper proposes a pipeline internal inspection intelligent diagnostic framework based on a hierarchical reward mechanism. This framework collects experiential data through a specific human-machine interaction mechanism and applies corresponding human-like reinforcement learning to achieve accurate and automated detection of defects and pipeline structures in pull-in pipelines. Subsequently, the framework is further refined and optimized using a hierarchical exploration mechanism to achieve relatively accurate and automated detection of internal defects and pipeline structures in field pipelines. The following sections will provide detailed explanations of the framework construction and framework optimization.

Fig. 4. Application environments of the pipeline internal inspection system. (a) Pull-in pipeline. (b) Field pipeline.

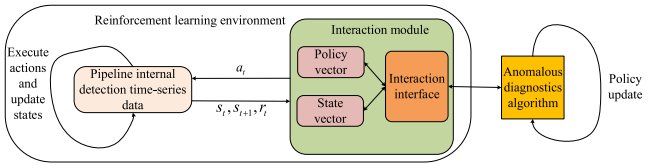


Fig. 5. Reinforcement learning environment for pipeline internal inspection intelligent diagnostic framework.

3.1. Construction of pipeline intelligent diagnostic framework

3.1.1. Parameter space modeling

The reinforcement learning process is a typical Markov decision process, and its corresponding parameter space can be represented as a quintuple (S,A,P,R,γ). In this paper, the reinforcement learning environment consists of the above quintuple, aiming to associate the pipeline eddy current data with the anomaly diagnostic algorithm and provide the operating environment for subsequent algorithm training. The training process of the entire anomaly diagnostics algorithm is as follows: Assuming the current time step is t, the reinforcement learning environment first determines the current state s_t based on the observed pipeline eddy current data and passes it to the diagnostics algorithm through the interaction interface. The algorithm makes a decision based on the current state and determines the optimal action a_t, which is then executed in the reinforcement learning environment. The corresponding pipeline eddy current data in the environment will be updated to the next time step state s_{t+1} under the influence of the optimal action a_t, and the reward function value r_t corresponding to the current action a_t and the updated next time step state s_{t+1} are fed back to the algorithm through the interaction interface, providing a basis for subsequent decision-making. Through the above steps, the reinforcement learning environment and the anomaly diagnostics algorithm form a complete closed-loop cycle, as shown below.

As the research objective of this paper is to achieve fast and accurate detection of anomaly regions and identification of anomaly types in complex pipeline eddy current data under different environments, it is essential to define and design the quintuple parameter space in a reasonable manner for subsequent discussions. The specific design process is supplied in the appendix.

The final design results of quintuple parameter space are shown in Table 1.

3.1.2. Action reward function derivation

During the interaction process, the first action taken by the agent in response to the pipeline eddy current signals is action a_1, which is a key action in this paper. The quality of window segmentation significantly affects the judgment results of subsequent actions. In other words, if action a_1 can accurately divide the pipeline eddy current signals into complete temporal signal segments (where each individual abnormal signal is fully contained within a single adaptive window), the agent can then perform fast and accurate attribute judgment (a_2) and type judgment (a_3) for different temporal signal segments. Therefore, designing a reasonable reward function to achieve the above objectives is the key problem to be addressed in this section.

Table 1
Design of specific action space.

Symbol of actions	Definition of actions	Range of actions
a_1	Adaptive window signal division action	{1, 2, 3, …, 1022, 1023, 1024}
a_2	Adaptive window attribute judgment action	{normal window, abnormal window}
a_3	Adaptive window type judgment action	{normal, bend, weld, defect}
a_4	The reliability of attribute judgment action	{0, 0.1, 0.2, …, 0.8, 0.9, 1.0}
a_5	The reliability of type judgment action	{0, 0.1, 0.2, …, 0.8, 0.9, 1.0}

Based on whether the adaptive window contains abnormal signals, the above key problem can be divided into two sub-problems:

1) If the adaptive window contains abnormal signals, the window should fully encompass the abnormal signals while keeping the window size as small as possible.
2) If the adaptive window does not contain any abnormal signals, under certain conditions, the window should be as large as possible.

Prior to reward function design, some parameter definitions are given as follows: Given the eddy current signals corresponding to the pull-in pipeline, the data points can be classified into four categories: normal, defect, weld, and bend, which correspond to label values 0, 1, 2, and 3, respectively. Let the current state be s, and the adaptive window action taken at state s corresponds to the data point range (m, n) where $n > m$. Let i represents any point within the adaptive window, and $label(i)$ denotes the corresponding label data for that point. In addition, let (m_j, n_j) denote the label range region of the j-th abnormal signal within the adaptive window.

For sub-problem 1), the reward function is designed as follows:

$$R_{11}^{*}(s,a_1)=\frac{p_{valid}-\alpha}{1-\alpha}\times r_{11}-\sum_{i=1}^{2}\mathrm{sgn}(label(k_i))\times r_{12}-R_{13}(s,a_1) \tag{1}$$

$$p_{valid}=\frac{\sum_{i=m}^{n}\mathrm{sgn}(label(i))}{n-m+1} \tag{2}$$

$$R_{13}(s,a_1) = -\operatorname{sgn}(count(a_1)-1)\times count(a_1)\times r_{13} \tag{3}$$

In Eqn. (1), p_{valid} represents the proportion of the abnormal signal's length within the corresponding adaptive window. Its calculation is described in Eqn. (2). α represents the threshold for determining whether p_{valid} receives positive or negative rewards. This threshold is determined based on subjective importance. k_1 and k_2 are two points, representing a proper and centered range in the adaptive window, where their values are determined subjectively and we choose 10%to 90% region ranges of the adaptive window here. r_{11} and r_{12} represent positive constant rewards. $R_{13}(s,a_1)$ is described in Eqn. (3), where *count* (a_1) represents the number of abnormal signals with intervals greater than the distance threshold parameter of 256 points, and r_{13} represents a positive constant reward.

According to Eqn. (1), the reward function for subproblem 1) consists of three parts: r_{11}, r_{12}, and $R_{13}(s,a_1)$. The r_{11} part measures the proportion of the abnormal signal segment within the adaptive window. When this proportion exceeds $\alpha(0 \le \alpha \le 1)$ threshold, a higher reward is given. Otherwise, if the proportion is smaller, a greater punishment is applied. The α threshold serves as the boundary for reward or punishment based on the proportion. The r_{12} part ensures abnormal signal properly and centrally contained within the adaptive window. When k_1 and k_2 represent normal data points, no punishment is incurred. However, if k_1 and k_2 are abnormal data points, a significant punishment is imposed, which overwhelms any rewards obtained from the r_{11} part. This encourages the agent to focus on the issue of whether the abnormal signal is located properly within the adaptive window. The $R_{13}(s,a_1)$ is used to assess whether there are multiple abnormal signals within the adaptive window, thereby promoting the accurate division of adaptive windows containing only a single abnormal signal by the agent.

For sub-problem 2), the reward function is designed as follows:

$$R^*_{21}(s,a_1) = \frac{n-m+1}{winSizeMax}\times r_{21} \tag{4}$$

where *winSizeMax* is the maximum window length that can be divided by the adaptive window action. r_{21} is a positive constant reward.

Therefore, the final reward function of action a_1 is shown below:

$$R^*_1(s,a_1) = \operatorname{sgn}(p_{valid})R^*_{11}(s,a_1) + (1-\operatorname{sgn}(p_{valid}))R^*_{21}(s,a_1) \tag{5}$$

After performing adaptive window action partitioning on the detected temporal signals inside the pipeline, multiple complete temporal signal segments are obtained. The next step is to make a thorough assessment of these signal segments, which involves two main actions: 1) determining the attribute of the temporal signal segment (normal or abnormal), denoted as action a_2, and 2) determining the type of the temporal signal segment (normal, defect, bend, or weld), denoted as action a_3. Since actions a_2 and a_3 are categorical and mutually exclusive, the reward function can be designed as follows: a fixed reward is given for correct judgments, while a fixed penalty is imposed for incorrect judgments. However, due to the inherent randomness in the judgment process (similar to random guessing), auxiliary actions a_4 and a_5 are introduced in this paper to mitigate such occurrences. The auxiliary actions essentially represent confidence indices, which measure the algorithm's level of certainty in correctly identifying a particular window signal and reflect the algorithm's mastery of different window signals. The key characteristic is that higher penalties are assigned when the diagnostic algorithm correctly identifies a window with low confidence or incorrectly identifies a window with high confidence. Conversely, when the diagnostic algorithm provides confidence indices that align with its true understanding, the rewards or penalties assigned to the identified windows are within a reasonable range. Therefore, a logarithmic form is used to design the reward function for the auxiliary actions.

The reward function for action a_2 is as follows:

$$R^*_{21}(a_2,a_4) = \begin{cases} \log_2(2a_4)\times r_{2i} & \text{correct answer} \\ -\log_2 2(1-a_4)\times r_{2i} & \text{false answer} \end{cases} \tag{6}$$

where the characteristics of this function are as follows: when a correct judgment is made with a confidence index of 100%, a full reward is obtained; when the confidence index is 50%, regardless of whether the judgment is correct or incorrect, no reward is given; when a correct judgment is made with a confidence index of 10% or an incorrect judgment is made with a confidence index of 90%, the obtained rewards are equivalent. These characteristics encourage the diagnostic algorithm to continuously strengthen its understanding of different window signals, aiming to achieve correct judgments with higher confidence indices [23].

Similarly, the reward function for action a_3 is as follows:

$$R^*_{22}(a_3,a_5) = \begin{cases} \log_4(4a_5)\times r_{2j} & \text{correct answer} \\ -\log_4\frac{4}{3}(1-a_5)\times r_{2j} & \text{false answer} \end{cases} \tag{7}$$

The reward functions for actions a_2 and a_3 can be combined into a single reward function to represent the overall performance of the window recognition actions. The function is defined as follows:

$$R^*_2(a_2,a_3,a_4,a_5) = q_{21}\times R^*_{21}(a_2,a_4) + q_{22}\times R^*_{22}(a_3,a_5) \tag{8}$$

where q_{21} and q_{22} are the proportion coefficients that measure the importance difference between action a_2 and action a_3.

Therefore, the reward function that measures the detection performance of the entire framework can be defined as follows:

$$R^*(s,a_1,a_2,a_3,a_4,a_5) = tR^*_1(s,a_1) + (1-t)R^*_2(a_2,a_3,a_4,a_5) \tag{9}$$

where $t \in [0,1]$ is an importance factor, and is set to a value of 0.5 to give equal importance to both R^*_1 and R^*_2.

3.1.3. Interactive process design

The intelligent diagnostic framework for pipeline internal detection based on a hierarchical reward mechanism is illustrated in Fig. 7 where reward function can be shown in Fig. 6. The specific interaction process is as follows: The environment consists of single-channel amplitude data of the pipeline's temporal detection signals, with the specific length being related to the actual length of the pipeline, typically in the order of tens of millions of data points. The agent represents an intelligent decision-making algorithm aimed at pipeline anomaly detection. At time t, the agent observes the current environment and its own current state s_t, and based on the internal decision-making algorithm, takes an action from the action set a_t. Subsequently, the action set a_t interacts with the environment, causing the current state s_t to transition to the next state s_{t+1}, and an immediate reward r_t is provided at time $t+1$ to evaluate the quality of the action a_t. The agent subsequently observes the state s_{t+1} and the reward r_t, where r_t, s_t, a_t, and other elements form a five-tuple $(s_t, a_t, r_t, P, \gamma)$ used for updating the decision-making algorithm. The state s_{t+1} then obtains a new action set a_{t+1} through the processing of the decision-making algorithm, and this loop continues until the exit condition is met. When the agent completes its training and is used to test

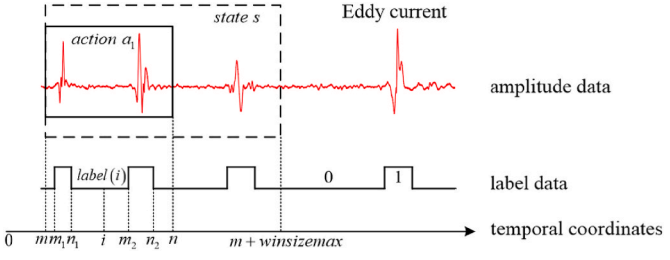


Fig. 6. Parameter definitions used in reward function design.

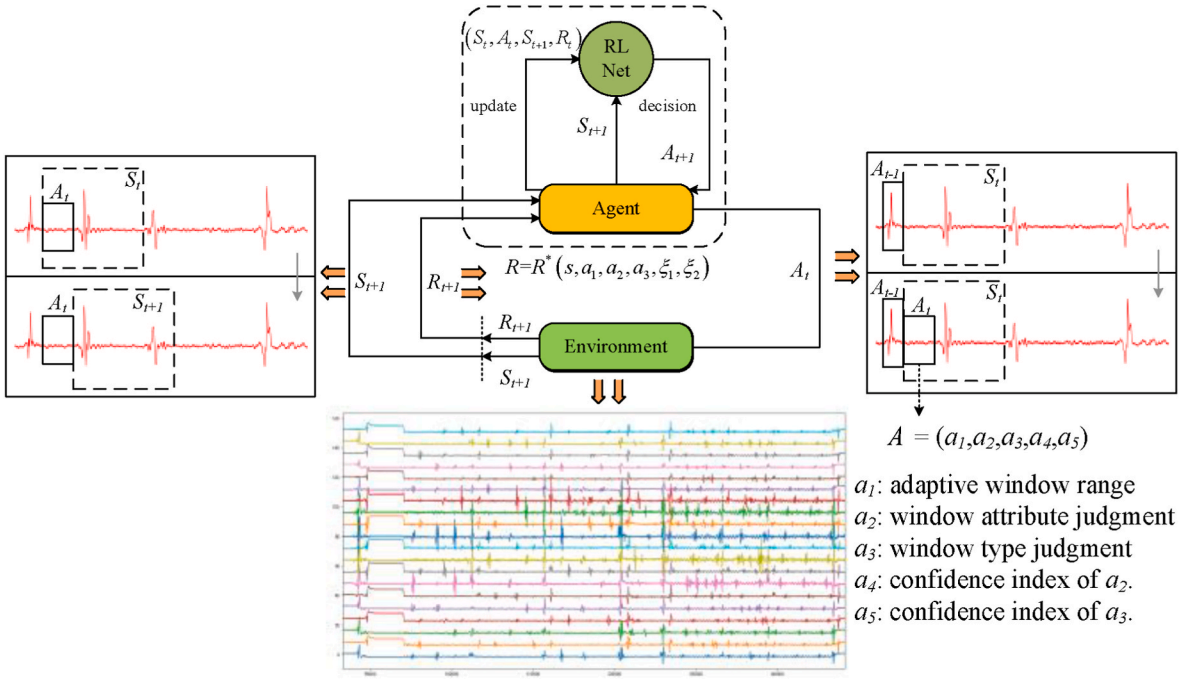


Fig. 7. Proposed framework for pipeline intelligent diagnostic system based on a hierarchical reward mechanism.

new temporal detection data from the pipeline, the interaction process remains the same as in Fig. 7, with the exclusion of the decision-making algorithm's decision update based on the five-tuple $(s_t, a_t, r_t, P, \gamma)$.

3.2. Optimization of pipeline intelligent diagnostic framework

The aforementioned intelligent diagnostic framework based on the hierarchical reward mechanism is mainly designed for tensioned pipeline data, but the final application scenario is field pipeline data during service. Compared to tensioned pipelines, field pipelines exhibit poorer internal detection environment, greater inherent signal noise, a wider range of pipeline states, and more complex characteristics of different defects. These factors result in lower quality of raw defect data and increased difficulty in subsequent analysis, making it impossible to meet all the requirements of practical pipeline inspections. Directly applying the aforementioned intelligent diagnostic framework to the automated detection of field pipeline signals would lead to unsatisfactory performance in terms of algorithm generalization and robustness.

Therefore, in this section, considering the characteristics of field pipeline signals, we propose to introduce a hierarchical exploration mechanism on top of the existing framework to achieve differentiated learning of raw data with different characteristics in field pipelines. This approach aims to enhance the generalization and robustness of the entire intelligent diagnostic framework to different field pipeline signals. The hierarchical exploration mechanism consists of two parts: temporal exploration and spatial exploration. The temporal exploration mechanism focuses on targeted exploration of different segments of single-channel raw eddy current data to enhance the framework's understanding of complex defect segments. The spatial exploration mechanism adaptively denoises different regions of multi-channel raw eddy current data to improve the framework's understanding of signal regions with severe noise interference.

The following provides a description of the optimization parts of the proposed framework.

3.2.1. Optimization to reward function

For field pipeline signals, due to the difficulties and high costs associated with on-site excavation work, they are often considered unlabeled data. To address the issue of ineffective action reward functions caused by the lack of labeled data for field pipelines, we introduce the use of power spectral features as a substitute for labeled data in the decision-making process. Therefore, the optimal reward function for the action a_1 is improved as follows. For subproblem 1), the optimized reward function is shown below.

$$R_{11}^{*}(s,a_1) = \frac{p_{valid} - \alpha}{1-\alpha} \times r_{11}^{'} - f(tp_1, tp_2) \times r_{12}^{'} \tag{10}$$

$$p_{valid} = \frac{tp_2 - tp_1 + 1}{n - m + 1} \tag{11}$$

$$f(tp_1, tp_2) = \begin{cases} 0 & tp_1 > k_1 \ and \ tp_2 < k_2 \\ 1 & others \\ 2 & tp_1 < k_1 \ and \ tp_2 > k_2 \end{cases} \tag{12}$$

where k_1, k_2 are the points within a certain distance from the left and right ends of the adaptive window, and tp_1, tp_2 are two points determined based on the power spectrum characteristics inside the adaptive window. Compared with Eqn. (1), Eqn. (10) uses $f(tp_1,tp_2)$ for r_{12} part calculation, and omits the penalty function $R_{11}(s,a_1)$ based on the number of abnormal signals within the adaptive window. This is because in field pipeline signals, closely spaced abnormal signals often couple together to form a complex "larger" abnormal signal, making it difficult to effectively differentiate them. In such cases, treating them as a single abnormal signal is sufficient.

For subproblem 2), due to the reward function is not connected to label, so it keeps unchanged shown in Eqn. (4). Therefore, the reward function of action a_1 is optimized as:

$$R_1^{*}(s,a_1) = \mathrm{sgn}(pe_{\max} - pe_{\mathrm{thre}}) R_{11}^{*} + (1 - \mathrm{sgn}(pe_{\max} - pe_{\mathrm{thre}})) R_{21}^{*} \tag{13}$$

where pe_{max} represents the maximum value of the power spectral density corresponding to the amplitude signal in the current adaptive window. pe_{thre} represents the adaptive parameter threshold of the power spectral density corresponding to the amplitude signal in the current pipeline section. Its actual meaning is the upper bound of the power spectral density for normal window signals, and its value is related to the overall complexity of the current pipeline section's signal (i.e., when the signal

noise is small or there are fewer overall anomalies, the value of pe_{thre} will adaptively decrease, and vice versa). In simple terms, pe_{thre} will adaptively adjust based on the overall signal conditions of different pipeline sections, and the region in the adaptive window corresponding to the power spectral density exceeding the pe_{thre} threshold will be considered as the true anomaly region.

For action a_2 or a_3, since it is a categorical decision, the derivation process of the optimal reward function remains unchanged, as shown above, with only the source of ground truth answers being modified. Due to the limitations of power spectral features, we can only determine the presence of abnormal signals for action a_2 and are unable to identify the specific signal types for action a_3, so the reward for a_3 is disregarded and the corresponding weighting coefficients for the reward functions are set to zero. Besides, the determination of specific signal types in field pipelines will be based on a new mechanism, as outlined in the following section.

The optimized reward function for action a_2 is shown below.

$$R_{21}^{*}(a_2,a_4)=\begin{cases}\log_2(2a_4)\times r_{21}^{'} & \text{anomaly and } pe_{\max}>pe_{\text{thre}}\\ \log_2(2a_4)\times r_{22}^{'} & \text{normal and } pe_{\max}\le pe_{\text{thre}}\\ -\log_2 2(1-a_4)\times r_{21}^{'} & \text{anomaly but } pe_{\max}\le pe_{\text{thre}}\\ -\log_2 2(1-a_4)\times r_{22}^{'} & \text{normal but } pe_{\max}>pe_{\text{thre}}\end{cases} \tag{14}$$

Since the main forms of the optimal reward functions for a_1 to a_5 have not changed, the fusion method of the functions remains the same, which is shown in Eqn. (9).

3.2.2. Design of temporal and spatial exploration mechanism

Before designing the exploration mechanism, we unify some relevant conceptual terms as shown in Fig. 8.

Compared to pull-in pipeline data, field pipeline data is more complex, primarily due to the following reasons:

1) Significant variation in signal noise among different pipeline sections. Some sections may experience significant noise interference due to sensor effects (such as equipment vibration, foreign object obstruction inside the pipe, or interference from other power frequency signals), while other sections have relatively low levels of noise.
2) Substantial variation in signal complexity among different pipeline sections. Some sections may contain numerous defects or spiral welds, resulting in highly intricate signal patterns, while other sections may have only a few defect signals, exhibiting simpler patterns.

To address these complexities, differential exploration of data features is proposed from both the temporal and spatial perspectives.

The basic process of the spatiotemporal exploration mechanism is illustrated in Fig. 9. The blue line represents the principle of the temporal exploration mechanism, which considers the lateral complexity of individual channel signals (i.e., comparing the lateral complexity of signals from different pipeline sections) to determine which individual channel signals require further attention and learning. The red line

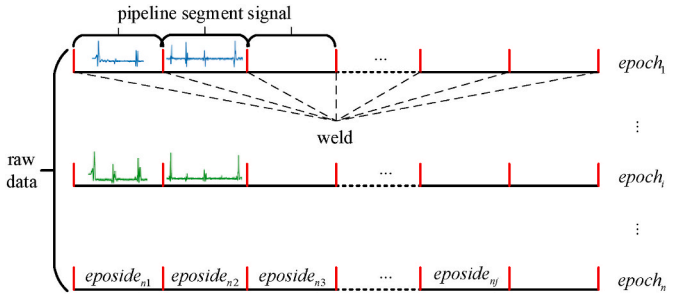


Fig. 8. Illustration of some conceptual terms used in exploration mechanism.

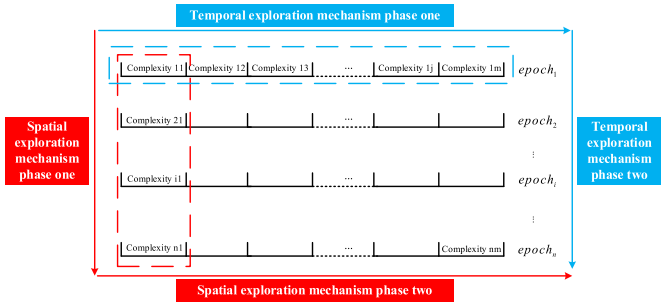


Fig. 9. Principle of temporal and spatial exploration mechanism.

represents the principle of the spatial exploration mechanism, which considers the vertical complexity of multi-channel signals (i.e., comparing the vertical complexity of signals from different pipeline sections) to determine which multi-channel signals require further attention and learning.

In relative terms, the temporal exploration mechanism partially addresses the issue of learning imbalance caused by uneven distribution of raw eddy current signals in field pipelines. It enables the pipeline inspection interaction model to perform targeted and differential learning on the raw eddy current signals from different pipeline sections, thereby improving the model's understanding of signals with different patterns. On the other hand, the spatial exploration mechanism targets pipeline sections with significant noise interference and effectively enhances the diagnostic algorithm's understanding of complex eddy current signals through directed denoising.

In addition, the basic feature of the spatiotemporal exploration mechanism, which is the complexity of the pipeline section signals, mainly includes two aspects: the level of noise interference and the variation in the distribution of anomalies. In this study, the number of identified abnormal signals obtained from the existing intelligent diagnostic framework is chosen as the measure. The reasons are as follows: When the overall noise of a pipeline section increases, the corresponding adaptive parameter threshold pe_{thr} also increases. Similarly, the power spectrum pe of each adaptive window signal within the section also increases. However, the growth trend of pe is much larger than that of pe_{thr}. This leads to many normal signal windows being misclassified as abnormal, resulting in an increased number of identified abnormal signal windows within the section. When the noise is severe, compared to other pipeline sections, the total number of identified abnormal signals across all channels in that section will be very large, thereby achieving the purpose of distinguishing different levels of noise interference. Furthermore, the number of identified abnormal signals in different pipeline sections can effectively characterize the differences in anomaly distribution. Therefore, this approach is reasonable.

3.2.3. Improvement of interactive process

The interactive process is improved as follows, where ①② represent the pipeline intelligent diagnostic framework based on the hierarchical reward mechanism, and ③④ represent the spatiotemporal exploration mechanism introduced in the optimization phase. The pipeline intelligent diagnostic framework optimized based on the hierarchical exploration mechanism is denoted by ①②③④.

As shown in Fig. 10, the optimized interactive process during the training phase is as follows:

1) For the pull-in pipeline data, the pipeline inspection intelligent diagnostic framework based on the hierarchical reward mechanism is used to train the intelligent decision algorithm and obtain the model pre-parameters.
2) For the field pipeline data, based on the model pre-parameters, the relevant data is trained using the optimized pipeline inspection

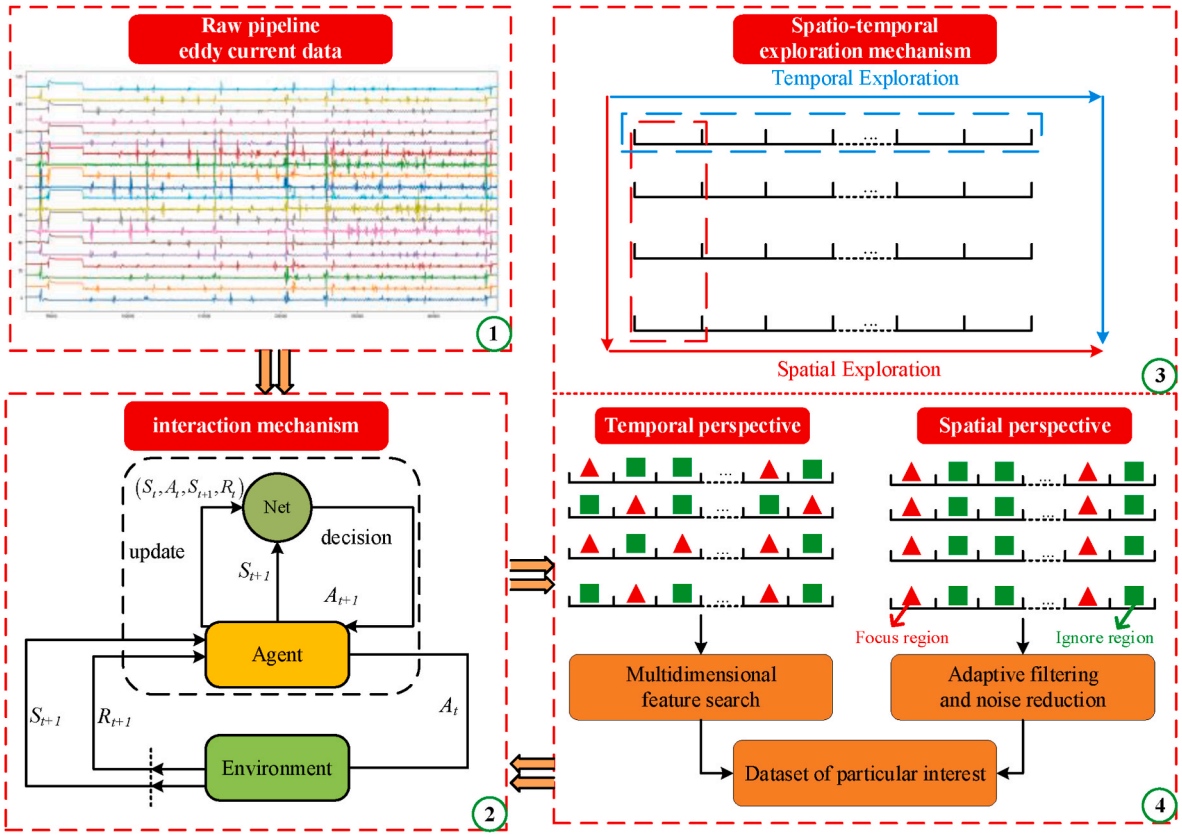


Fig. 10. Proposed framework for pipeline intelligent diagnostic system after the optimization of temporal and spatial exploration mechanism.

intelligent diagnostic framework to obtain the final model parameters.

In this way, the framework is capable of learning various features of field pipeline eddy current data based on the knowledge gained from the pull-in pipeline, which enhances the cognitive ability of the framework in recognizing complex signals in the pipeline, making it possible to realize the automated detection of field pipeline.

4. Experiment and result analysis

4.1. Model indexes and experimental data

4.1.1. Model indexes

For the pipeline internal defect anomaly diagnostics problem studied in this paper, due to the rarity of abnormal signals within the entire signal, the introduction of metrics such as precision and recall is necessary to evaluate the model's performance. Precision represents the proportion of true positive predictions among the predicted positive samples, while recall represents the proportion of true positive samples correctly predicted as positive by the model. Typically, these two metrics are in conflict with each other, meaning that while pursuing improvement in one metric, the other metric may be negatively affected.

For binary classification problems, their states can be divided into four categories: true positives (*TP*), true negatives (*TN*), false positives (*FP*), and false negatives (*FN*), so the calculation methods for precision and recall are as follows:

$$precision = \frac{TP}{TP + FP} \tag{15}$$

$$recall = \frac{TP}{TP + FN} \tag{16}$$

When generalizing the problem to n-class classification, the calculation methods for precision and recall are as follows:

$$precision = \frac{TP_{11}}{TP_{11} + \sum_{k=2}^{n} FP_{1k}} \tag{17}$$

$$recall = \frac{TP_{11}}{TP_{11} + \sum_{k=2}^{n} FP_{k1}} \tag{18}$$

The definitions of TP_{nn} and FP_{nn} are shown in Fig. 11, where the matrix is referred to as the confusion matrix, serving to assess the overall classification performance of the model. When the data is distributed along the diagonal, the model exhibits the best performance in multi-class classification.

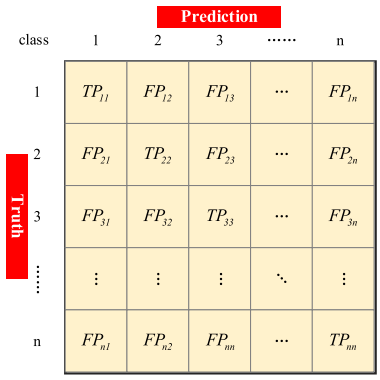


Fig. 11. Schematic diagram of confusion matrix of multi classification problem.

To comprehensively evaluate the model's classification performance, this paper introduces the F-score, which combines precision and recall. The calculation method is as follows:

$$F-score=\left(1+\beta_f{}^2\right)\times\frac{precision\times recall}{\left(\beta_f{}^2\times precision\right)+recall} \tag{19}$$

Considering the ultimate application scenario of pipeline internal defect anomaly detection in this paper, where the priority is given to recall to identify all true defects as much as possible, we set the value of β_f as 2.

4.1.2. Experimental data

The experimental data in this paper is divided into two categories: pull-in pipelines and field pipelines. The pull-in pipelines are classified based on their diameter sizes into three types: 168 pipeline, 219 pipeline, and 273 pipeline (with a diameter of 273 mm). Among them, partial true value parameters for defects in the 219 pipeline are shown in Table 2, and more related true value information for defects can be found in the supplementary materials.

In comparison to the pull-in pipelines where accurate true value defect parameters were obtained through manual carving, the field pipelines completely lack such data due to difficulties in construction and other engineering reasons. To facilitate subsequent experimental discussions, alternative methods were employed to obtain relatively accurate reference defect values for the field pipelines. Specifically, the positions of pipeline structural components were derived from the initial design drawings of the pipelines and manually annotated, while the positions of pipeline defects were obtained from the corresponding detection reports generated by a state-of-the-art commercial algorithm used in the field. Additionally, some positions were further annotated by combining with partial manual excavation verification processes.

4.2. Experiments on the unoptimized pipeline intelligent diagnostic framework

4.2.1. Effects of different actions

Due to the pipeline internal detection intelligent diagnostic framework involving three intermediate steps for time-series eddy current data, namely, adaptive window signal segmentation, adaptive window attribute classification, and adaptive window type classification, the actual performance of these three steps will be analyzed.

(a) Adaptive window signal segmentation action

Regarding the adaptive window signal segmentation, the concept of a confusion matrix is introduced, and the window types are classified into three categories: A, B, and C. Category A represents the case where a single anomaly signal is accurately segmented into a single adaptive window; category B represents the case where a single anomaly signal is segmented into multiple adaptive windows; and category C represents the case where multiple anomaly signals are segmented into a single adaptive window. The specific results of the confusion matrix are shown in Fig. 12. Since the actual counts for categories B and C should be zero from a true classification perspective, all the second and third rows of the confusion matrix are marked as 0. Subsequently, the analysis will focus on the data in the first row of the confusion matrix.

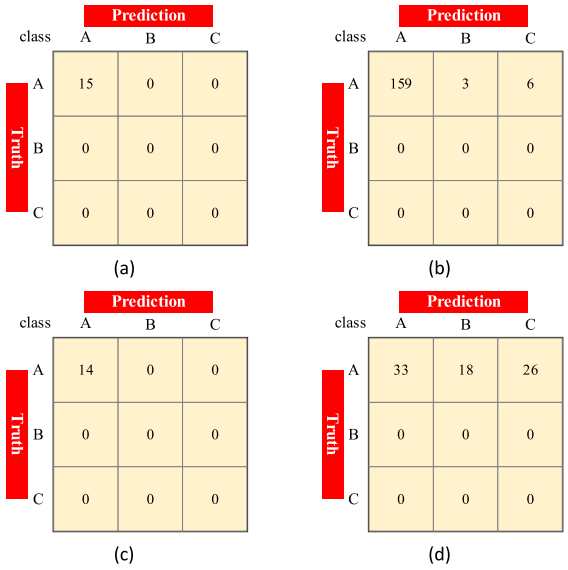


Fig. 12. Confusion matrix of adaptive window signal segmentation action. (a) Weld signals in pull-in pipeline. (b) Defect signals in pull-in pipeline. (c) Weld signals in field pipeline. (d) Defect signals in field pipeline.

From Table 3, it can be observed that there are a total of 15 weld signals and 168 defect signals in the pull-in pipelines, while there are 14 weld signals and 77 defect signals in the field pipelines. In Fig. 12(a) and (b), the window segmentation action achieves a recall rate of 100% for weld signals in the pull-in pipelines (predicted as category A: 15/15), and a recall rate of 94.6% for defect signals (predicted as category A: 159/168). However, a few defect signals are misclassified as category B (3/168) or category C (6/168). This may be due to the presence of helical weld sections in the pull-in pipelines, where the signal characteristics are closer to those of field pipelines, resulting in more complex features compared to other pull-in sections and leading to misclassification. In Fig. 12(c) and (d), the window segmentation action achieves a high recall rate of 100% for weld signals in the field pipelines (predicted as category A: 14/14). However, the recall rate for defect signals is only 42.9% (predicted as category A: 33/77), with a large number of defect signals misclassified as category B (18/77) or category C (26/77). This is attributed to the fact that the field pipelines have poorer internal environment, larger original signal noise, and more complex defect morphologies, making it challenging for the current intelligent diagnostic framework to fully learn the effective features of relevant defect signals.

Therefore, the conclusions of window segmentation action on the raw data of oil and gas pipelines are as follows: (1) For pull-in pipelines, the window segmentation action demonstrates excellent partitioning capability for all abnormal signals (regardless of weld or defect signals). It accurately and completely segments each individual abnormal signal into a single adaptive window. (2) For field pipelines, the window segmentation action exhibits excellent partitioning capability for weld-related abnormal signals, accurately segmenting them into individual windows. However, it shows poorer performance for defect-related abnormal signals. Some individual defect-related abnormal signals are

Table 2
Truth table of some defects in pull-in 219 pipeline.

Defect Index	Length (mm)	Width (mm)	Depth (mm)	Distance to weld(m)	Clock position (°)
W1-1	56	53	4.27	1.60	123.8
W1-2	53	37	2.43	1.59	82.5
W1-3	55	55	3.87	1.60	15.5
W1-4	51	31	3.27	1.60	319.8
W1-5	51	44	3.85	1.60	247.6
W1-6	60	53	2.86	1.15	294.0

Table 3
Truth numbers of anomaly in related pull-in pipelines.

	Defect numbers	Weld numbers	Bend numbers
Pull-in 168 pipeline	43	5	0
Pull-in 219 pipeline	68	5	0
Pull-in 273 pipeline	55	5	0
Zhejiang Field pipeline	77	14	0

partitioned into multiple adaptive windows, while in other cases, multiple defect-related abnormal signals are partitioned into a single adaptive window.

(b) Adaptive window attribute judgment action

Regarding the segmented window set, ensuring the accurate identification of individual windows containing abnormal signals essentially constitutes a binary classification problem. The specific identification results are illustrated as follows.

From Table 4 and Table 5, it is evident that without confidence indices, the recall rate for defect-type windows in pull-in pipelines is 60.7% (102/168), and in field pipelines, it is 40.3% (31/77). However, with confidence indices, the recall rate for defect-type windows in pull-in pipelines increases to 100% (168/168), and in field pipelines, it rises to 66.2% (51/77). This indicates that the inclusion of confidence indices greatly improves the recognition performance of the attribute classification action for the abnormal states of adaptive windows.

Furthermore, from Table 6, with confidence indices, the recall rate for weld signals in pull-in pipelines is 100%, and the precision rate is 100%, while for defect signals, the recall rate is 100%, and the precision rate is 98.9%. In field pipelines, the recall rate for weld signals is 100%, and the precision rate is 100%, while for defect signals, the recall rate is 66.2%, and the precision rate is 73.9%.

Comparatively, it can be concluded that with confidence indices, the attribute classification action has the following actual recognition performance for the abnormal states of adaptive windows: (1) The attribute classification action demonstrates excellent recognition ability for both weld and defect signals in pull-in pipelines, with recall and precision rates at around 99%. (2) For field pipelines, the attribute classification action also exhibits excellent recognition ability for windows containing weld signals, with recall and precision rates at 100%. However, for windows containing defect signals, it shows a poorer recognition performance with a recall rate of only around 70%. The reasons are twofold: (a) Some defect signals are submerged amidst the complex noise interference in field pipelines, leading to misclassification as normal windows, resulting in a lower recall rate for defect windows. (b) Some defect signals are segmented into multiple adaptive windows, leading to the classification of multiple abnormal windows by the attribute classification action, resulting in a lower precision rate for defect windows.

(c) Adaptive window type judgment action

Similarly, for the window type classification (normal, weld, elbow, defect), it belongs to a typical multi-classification problem. Therefore, the effectiveness is directly measured using a confusion matrix, and the specific results are shown in Fig. 13.

From Fig. 13(a), it can be observed that the adaptive window type classification action exhibits excellent classification performance for abnormal signal categories in pull-in pipelines, with a recall rate of 95.2% for defect signals (160/168) and 93.3% for weld signals (14/15). However, there are still some shortcomings, such as 12 normal signals being misclassified as defect signals, which could be due to interference from spiral segment data in the pull-in pipelines with similar signal characteristics to field signals. Additionally, 8 defect signals are misclassified as weld signals, possibly due to the large amplitude of these defect signals, making them appear similar to weld signals in amplitude-based classification. Furthermore, 1 weld signal is misclassified as a defect signal, likely for a similar reason, where this specific type of weld signal has smaller amplitude features, resembling defect signals and leading to misclassification.

Table 4
Classification performance of the adaptive window attribute judgment action without confidence index.

Without confidence index							
Pull-in pipeline				Field pipeline			
Weld types		Defect types		Wele types		Defect types	
p	*r*	*p*	*r*	*p*	*r*	*p*	*r*
15/15	15/15	102/102	102/168	14/14	14/14	31/43	31/77

Table 5
Classification performance of the adaptive window attribute judgment action with confidence index.

With confidence index							
Pull-in pipeline				Field pipeline			
Weld types		Defect types		Wele types		Defect types	
p	*r*	*p*	*r*	*p*	*r*	*p*	*r*
15/15	15/15	168/170	168/168	14/14	14/14	51/69	51/77

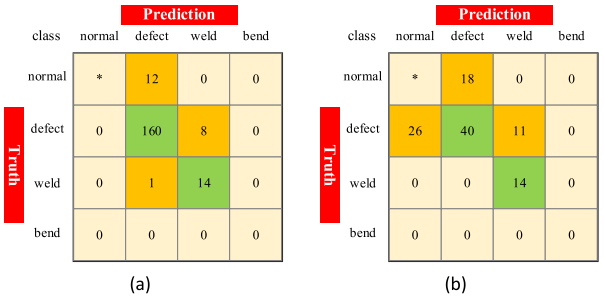


Fig. 13. Confusion matrix of the adaptive window type judgment action. (a) Pull-in pipeline. (b) Field pipeline.

Similarly analyzing Fig. 13(b), it can be observed that the adaptive window type classification action achieves excellent actual classification performance for abnormal signals in field pipelines, with a recall rate of 100% for weld signals (14/14). However, the recall rate for defect signals is only 51.9% (40/77), and the reasons behind this are twofold: (a) Severe noise interference in the field pipelines causes defect signals with relatively small amplitude changes to be submerged and misclassified as normal signals (26/77); (b) Severe noise interference in the field pipelines causes defect signals with relatively large amplitude changes to be misclassified as weld signals (11/77). In addition, since neither the pull-in pipelines nor the field pipelines' data in this experiment contain elbow signal information, the relevant positions in the confusion matrix are all marked as 0.

In summary, the actual classification performance of the adaptive window type classification action for abnormal signals can be concluded as follows: (1) For pull-in pipelines, this action can accurately identify and classify weld signals and defect signals, with recall rates around 95%. (2) For field pipelines, this action can accurately classify weld signals, maintaining a recall rate of over 95%. However, the performance is poor for defect signals, with a recall rate of only 52%.

4.2.2. Comparison with other methods

Due to the proprietary nature of most algorithms in the pipeline inspection field, where the source code or related information is not open-source or publicly available, it is not feasible to conduct a direct comparison of actual performance. Therefore, in this study, a comparison is made with multi-classification algorithms.

The pipeline status is divided into four categories: normal, defect, weld, and elbow. Traditional machine learning algorithms and deep learning-related algorithms are tested separately on pull-in pipelines (including three types of pipelines: 168, 219 and 273 pipelines). The traditional machine learning algorithms include K-Nearest Neighbor (KNN), Classification and Regression Tree (CART), Local Outlier Factor

(LOF), and Isolation Forest (iForest). The deep learning-related algorithms include Multi-Layer Perceptron (MLP) [24], Long Short Term Memory-Fully Convolutional Networks (LSTM-FCN) [25], and Multi-Scale Convolutional Neural Networks (MS-CNN) [26]. Since pull-in pipelines do not have elbow information, the parameters related to elbows are marked as 0. The specific test results are shown in Table 6.

However, since the ultimate application scenario of pipeline internal detection is the field environment, although ours performs extremely well in pull-in pipelines, as mentioned earlier, its performance in real field pipelines is not satisfactory. Specifically, the overall performance is: weld precision rate is above 95%, but the defect precision rate is only 52%; this is manifested as follows: (1) the ability of the adaptive window signal segmentation action to accurately partition individual abnormal signals into single windows needs improvement; (2) the recognition ability of the adaptive window signal identification action for window signals with severe noise interference needs improvement. Based on the above analysis, the fundamental reason for the poor performance lies in the fact that compared to pull-in pipe signals, real field pipelines suffer from poorer internal inspection environment, larger original signal noise, a greater variety of pipeline conditions, and more complex defect characteristics. These factors result in poor quality of original defect data, leading to greater difficulty in subsequent signal analysis. Therefore, these issues will be addressed by optimizing the framework.

4.3. Experiments on the optimized pipeline intelligent diagnostic framework

4.3.1. Ablation experiments on exploration mechanism

To test the impact of spatiotemporal exploration mechanisms on the automated detection capability of pipeline internal defects in real-world environments, this subsection validates the actual effects of the exploration mechanisms in three directions based on two types of frameworks: the unoptimized and optimized frameworks. The main comparisons include: (1) a comparison of the effectiveness of the adaptive window signal segmentation action, (2) a comparison of the effectiveness of the adaptive window attribute judgment action, and (3) a comparison of the effectiveness of the adaptive window type judgment action.

(a) Adaptive window signal segmentation action

Regarding the adaptive window signal segmentation action, this section conducts a two-step analysis: (1) a visual comparison of pipeline signal window partitioning, and (2) a comparison of pipeline signal window partitioning metrics.

For (1), specific comparisons are shown in Fig. 14 (supplied in appendix).

For (2), we conduct an evaluation method similar to the one used in Section IV B, the results are shown in Fig. 14.

From Fig. 14(a) and (b), it is evident that for the traction pipeline, the intelligent diagnostic framework without the spatiotemporal exploration mechanism exhibits inaccuracies and inadequacies in the window partition of defect signals (as indicated by classes B and C in the figure), which can be attributed to the presence of spiral pipeline signals. However, with the introduction of the spatiotemporal exploration mechanism, the optimized intelligent diagnostic framework demonstrates accurate identification and window partitioning of such spiral signals, resulting in a reduction of incorrect window partitions (classes B and C) to zero and an overall improvement in the window partition accuracy by 5.4% (from 94.6% to 100%).

From Fig. 14(c) and (d), it is evident that for the window partition of defect signals in the field pipeline, the introduction of the spatiotemporal exploration mechanism significantly enhances the intelligent diagnostic framework's ability to recognize defect signals, leading to a substantial improvement in the window partition accuracy from 42.9% to 92.2%, representing an overall increase of approximately 50%.

Based on the analyses, we draw the following conclusion: The spatiotemporal exploration mechanism can significantly enhance the adaptive window signal segmentation action's ability to recognize and

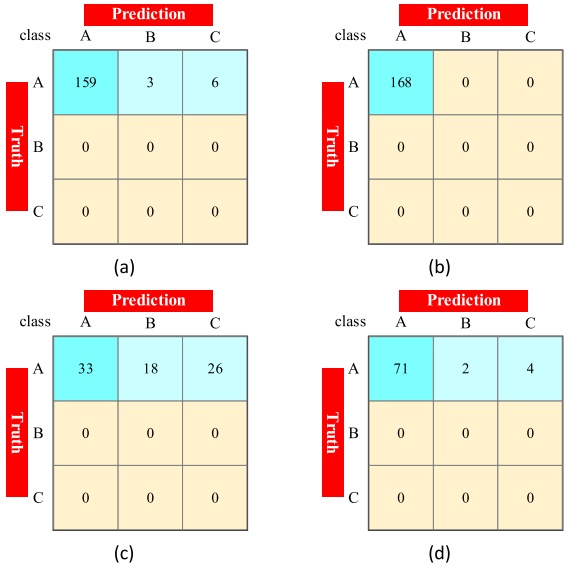


Fig. 14. Confusion matrix of the adaptive window signal segmentation action (a) Pull-in pipeline without exploration mechanism. (b) Pull-in pipeline with exploration mechanism. (c) Field pipeline without exploration mechanism. (d) Field pipeline with exploration mechanism.

Table 6
Comparison of different multi-classification algorithms based on pull-in pipelines.

Pull-in pipeline state types	Traditional machine learning field											
	KNN			CART			LOF			iForest		
	p	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*
normal	1.00	0.99	0.99	0.99	1.00	0.99	0.97	0.98	0.98	0.99	1.00	0.99
bend	0	0	0	0	0	0	0	0	0	0	0	0
weld	1.00	0.93	0.99	0.92	0.73	0.76	0.88	0.93	0.92	1.00	0.80	0.83
defect	0.57	0.73	0.69	0.73	0.43	0.47	0.79	0.64	0.67	1.00	0.61	0.66
Pull-in pipeline state types	Deep learning related fields											
	MLP			LSTM-FCN			MS-CNN			**Ours**		
	p	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*
normal	0.97	0.84	0.86	0.98	0.85	0.87	1.00	0.95	0.96	**1.00**	**0.99**	**0.99**
bend	0	0	0	0	0	0	0	0	0	0	0	0
weld	1.00	0.93	0.94	0.91	0.66	0.70	0.92	0.80	0.82	**1.00**	**1.00**	**1.00**
defect	1.00	0.41	0.46	0.00	0.00	0.00	0.81	0.48	0.52	**0.93**	**1.00**	**0.99**

partition complex signals (such as spiral pipeline signals or field pipeline signals).

(b) Adaptive window attribute judgment action

Regarding the adaptive window attribute judgment action, the comparative results are as follows.

According to Table 7, the introduction of the spatio-temporal exploration mechanism has significantly improved the intelligent diagnostic framework's recall rate for defect-type windows in real-world pipelines, increasing it from 66.2% (51/77) to 97.4% (75/77), representing an overall increase of 31.2%. The precision rate also increased from 73.9% (51/69) to 84.3% (75/89), showing an overall increase of 10.4%.

Clearly, the optimized intelligent diagnostic framework demonstrates a significant enhancement in its ability to identify abnormal states of signal windows. There are two main reasons for this improvement: (1) The spatio-temporal exploration mechanism greatly improves the accuracy of adaptive window signal segmentation, making it easier to perform attribute judgments based on signal windows and avoiding misidentifications and omissions caused by inaccurate window segmentation. (2) The corresponding feature processing methods in the spatio-temporal exploration mechanism, such as multi-dimensional feature extraction and adaptive filtering for noise reduction, significantly enhance the adaptive window attribute judgment action's capability to recognize abnormal signals in complex conditions.

Therefore, we can conclude that the spatio-temporal exploration mechanism significantly improves the adaptive window attribute judgment action's ability to identify abnormal states within signal windows under complex conditions.

(c) Adaptive window type judgment action

Regarding the adaptive window type judgment action, the comparative results are as follows. It is important to note that the optimized framework does not include a direct adaptive window type judgment action for signal windows. Instead, it relies on a step-by-step classification process based on the combination of adaptive window signal segmentation and adaptive window attribute judgment actions, along with the utilization of global spatiotemporal features. The basic approach is as follows: (1) Utilize the adaptive window signal segmentation action and the adaptive window attribute judgment action to obtain all abnormal state windows. (2) Use multi-channel spatial features to determine signal windows belonging to the weld type. (3) Utilize spatiotemporal features between pipe sections to identify segments containing bends. (4) The remaining abnormal state windows are considered as defect-type windows.

Based on Fig. 15, it can be observed that the optimized framework shows significant improvements in recognizing normal, defect, and weld signals. Specifically: (1) The number of normal signals mistakenly classified as defects due to noise interference reduced from 18 to 4, and all four were caused by significant sensor jitter, resulting in local intense fluctuations, making their waveform similar to defects. (2) The correct recognition of defect signals increased from 40 to 75, and the number of misclassifications as normal or weld signals significantly decreased. (3)

Table 7
Classification effects of the adaptive window attribute judgment action on field pipelines.

Field pipeline							
Without exploration mechanism				With exploration mechanism			
Weld types		Defect types		Wele types		Defect types	
p	*r*	*p*	*r*	*p*	*r*	*p*	*r*
14/14	14/14	51/69	51/77	14/14	14/14	75/89	75/77

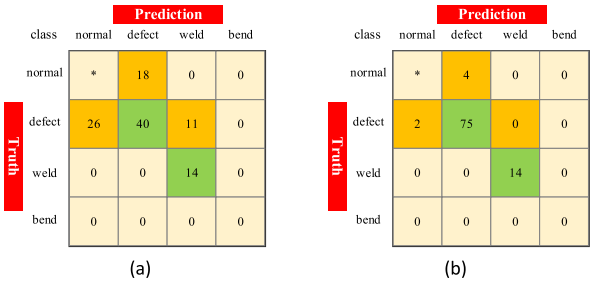


Fig. 15. Confusion matrix of the adaptive window type judgment action in field pipeline. (a) without exploration mechanism. (b) with exploration mechanism.

The number of misclassifications for weld signals significantly reduced, achieving 100% accurate recognition for them. Overall, the spatiotemporal exploration mechanism increased the defect signal recall rate from 51.9% to 97.4% and the precision rate from 69% to 94.9%; the weld signal recall rate remained unchanged at 100%, and the precision rate improved from 56% to 100%.

Therefore, the conclusion can be drawn that the spatiotemporal exploration mechanism significantly enhances the intelligent diagnostic framework's classification and recognition capabilities for abnormal state windows.

4.3.2. Verification of robustness and generalization

In the previous section, the effectiveness and necessity of the spatiotemporal exploration mechanism were validated through ablation experiments. In this section, the overall detection performance of the optimized intelligent diagnostic framework will be tested based on real-world pipeline detection results. This testing comprises two main parts: framework robustness validation and generalization validation.

(a) Algorithm comparison to validate robustness

Different algorithms are compared to assess the optimized framework's ability to perform well under various conditions and resist performance degradation due to changes in the operating environment. The specific results are as follows.

According to Table 8, it can be observed that, compared to the pull-in pipeline, the performance of all classifiers in the real-world field pipeline generally deteriorates. However, the optimized pipeline internal detection intelligent diagnostic framework shows a significant improvement in performance for the real-world field pipeline. Compared to the best-performing KNN classification algorithm for the real-world field pipeline, the optimized framework demonstrates superior recognition performance for both weld-type and defect-type signals, specifically: the weld recall rate increases from 0.89 to 0.93 (an improvement of 4%), the defect recall rate increases from 0.65 to 0.89 (an improvement of 24%), and the overall detection performance (measured by F-score) improves by 15%.

Therefore, the optimized intelligent diagnostic framework exhibits better robustness in detecting anomalies in real-world field pipelines.

(b) Pipeline comparison to validate generalization

The optimized framework will be tested on different pipelines to evaluate its ability to generalize and adapt to diverse pipeline structures and characteristics. Here, a real pipeline segment, approximately 4 km in length, from the Changqing Oilfield Eastern Mainline is chosen as the detection target. Parts of raw eddy current data consisting of pipeline structures(weld, bend) and pipeline defects are shown in Fig. 16.

It is important to note that all previous pipeline experiments involved the pipelines listed in Table 4, which had the characteristic of having a reference list of true defects for validation. However, the

Table 8
Comparison of different multi-classification algorithms based on field pipelines.

Pull-in pipeline state types	Traditional machine learning field											
	KNN			CART			LOF			iForest		
	p	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*
normal	0.98	0.99	0.99	0.96	0.99	0.98	0.96	0.98	0.98	0.98	0.99	0.99
bend	0	0	0	0	0	0	0	0	0	0	0	0
weld	0.87	0.89	0.89	0.83	0.71	0.73	0.88	0.79	0.81	0.63	0.86	0.80
defect	0.56	0.65	0.63	0.69	0.37	0.41	0.66	0.48	0.51	0.71	0.53	0.56
Pull-in pipeline state type	Deep learning related fields											
	MLP			LSTM-FCN			MS-CNN			**Ours**		
	p	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*	*p*	*r*	*F*
normal	0.96	0.79	0.82	0.98	0.82	0.85	0.99	0.91	0.92	0.99	0.99	0.99
bend	0	0	0	0	0	0	0	0	0	0	0	0
weld	0.75	0.64	0.66	0.62	0.57	0.58	0.85	0.79	0.80	**0.96**	**0.93**	**0.94**
defect	0.81	0.37	0.42	0.00	0.00	0.00	0.77	0.41	0.45	**0.86**	**0.89**	**0.88**

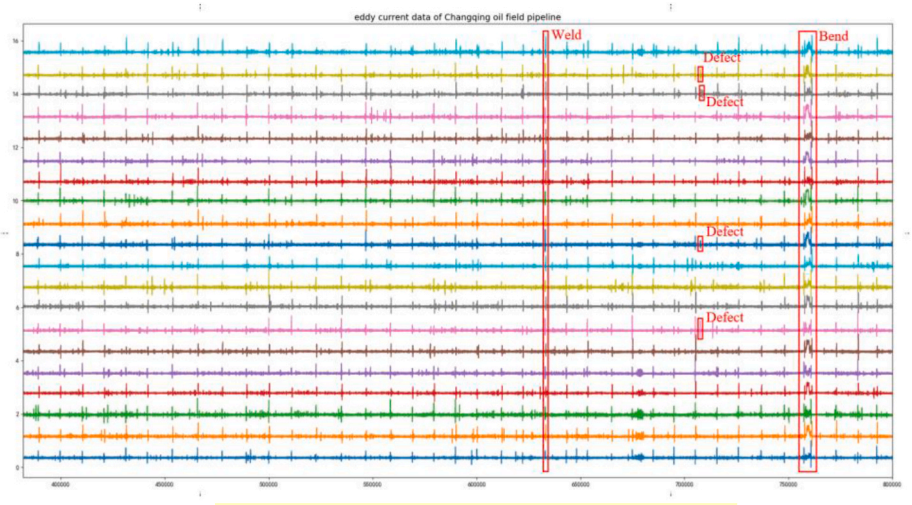


Fig. 16. Raw eddy current data of Changqing oil field pipeline.

chosen real pipeline in this instance does not have a reference list of true defects, so the actual detection results will be verified through manual comparison and validation. The specific detection results are shown in Table 9, where the actual quantities of welds and bends are manually counted based on the pipeline design drawings.

From Table 9, it can be observed that the optimized framework achieves a high detection rate for pipeline structural components such as welds and bends, with the detection rate for welds being over 95% and for bends being over 90%. However, regarding defect signals, the algorithm automatically detects a total of 1273 defects, and upon signal analysis, it is found that many of these detected defects are small defects (i.e., signals with relatively small amplitude variations), making it difficult to individually verify each defect through manual intervention. Therefore, to verify the results, reference was made to detection reports from multiple foreign companies for the same pipeline, and a total of 30 largest defects were manually selected for validation. The results indicate that the optimized framework successfully detected all 30 defect signals.

Furthermore, if the common anomalies (bends, welds, defects) from the above-mentioned company detection reports are taken as the ground truth reference, the corresponding confusion matrix can be obtained, as shown in Fig. 17.

The percentages in the matrix represent the precision rate, which indicates the proportion of anomalies detected by the optimized framework among the common anomalies, where it can be observed that the optimized framework exhibits excellent recognition performance for bends, welds, and defects, with precision rates of 92.3% for bends,

Table 9
Results of anomaly detection in Changqing oil field pipeline.

Anomaly types	Verification methods	Anomaly numbers	Detection numbers	Detection ratio
Weld	Algorithm	359	351	97.77%
Bend	Algorithm	52	48	92.31%
Defect	Manual	30	30	100%
	Algorithm	Difficult to calculate	1273	Difficult to verify

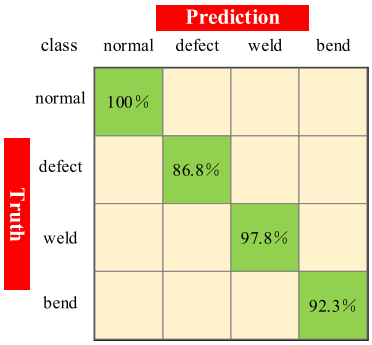


Fig. 17. Confusion matrix of anomaly detection in Changqing oil field pipeline.

97.8% for welds, and 86.8% for defects.

Therefore, based on the comprehensive analysis, it can be concluded that the optimized framework exhibits better generalization ability in the detection of anomalies in real-world field pipelines.

5. Conclusion

In response to the challenges of severe noise interference, complex anomaly states, and low automation in the detection of real-world pipeline signals, this paper has proposed a pipeline internal detection intelligent diagnostic framework based on reinforcement learning with hierarchical reward exploration mechanism to achieve accurate and automated detection of complex internal pipeline signals. This paper has constructed the entire pipeline internal detection intelligent diagnostic framework using the hierarchical reward mechanism. By adaptively dividing the original pipeline data into sets of signals with different window sizes and performing attribute and type judgments, the accurate and automated detection of defects in the pull-in pipeline is achieved. Secondly, the hierarchical exploration mechanism has been introduced to optimize the entire intelligent diagnostic framework. Through spatiotemporal dimensionality, deep search, and feature learning of complex signals in the real-world field pipeline, a relatively accurate automated detection of internal defects in the real-world field pipeline is achieved. Building upon this, the next step in the future will involve in-depth quantification of defects in the detected abnormal signal windows.

CRediT authorship contribution statement

Lanqin Su: Writing – original draft, Software, Methodology. **Bin Gao:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Xiangyu Zhao:** Software, Methodology, Data curation. **Yang Fu:** Validation, Software. **Wai Lok Woo:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

The work received support from the Deyuan and UESTC Joint Research Center, as well as the National Natural Science Foundation of China (No. 61971093 and No. 61527803). Additionally, it was funded by the International Science and Technology Innovation Cooperation Project of Sichuan Province under grant number 2021YFH0036, as well as the Science and Technology Department of Sichuan, China, through Grant No. 2018JY0655 and Grant No. 2018GZ0047.

Appendix

A. Pipeline environment difference

Table 1
Comparison of application environments for pipeline internal inspection systems.

Types	Pipeline conditions	Supplementary notes
Pull-in Pipeline	Length of pipeline is short	Limited by the experimental site, usually several hundred meters.
	Structure of pipeline is simple	Does not contain elbow structural parts, and the overall number is small.
	Inner environment is good	Not in service, free from various types of residues and relatively clean inside the pipeline.
	Flaws of pipeline is simple	Defect shape is clear and noise interference is small. Position of flaws is known and contains ground truth.
Field Pipeline	Length of pipeline is long	Actual production application, usually more than 10 km.
	Structure of pipeline is complex	Contains structural parts such as elbows, welds, flanges, etc., and the overall number is large
	Inner environment is terrible	Contains all kinds of tiny impurities remaining in the transmission medium during service
	Flaws of pipeline is complex	Complex defect shape and larger noise interference.
		Unknown position of flaws and without ground truth.

B. Parameter space modeling

Design of state space S: It should satisfy the following two conditions: 1) Sufficiently represent changes in the environmental state; 2) Reduce unnecessary redundant information. It is known that the pipeline eddy current signal is a sequential data with N channels, and each channel corresponds to a distinct angle of the circular pipelines and contains two dimensions: amplitude and phase. Since the signal amplitude can already significantly reflect the characteristics of anomaly regions in most cases, this paper adopts a fixed length of amplitude data points from a single channel for state design.

Design of action space A: It involves the introduction of five actions to achieve the desired objectives, defined as (a_1, a_2, a_3, a_4, a_5), as presented in Table 2. Among them, the purpose of action a_1 is to divide the pipeline eddy current signal into multiple adaptively sized window signal segments according to the specified requirements. These signal segments can be classified into two types: 1) adaptive window signals that contain only a single abnormal region, and 2) normal window signals that do not contain any abnormal regions. This action is a crucial step in the reinforcement learning model as the quality of signal segmentation directly affects the subsequent window attribute judgment (action a_2) and window type judgment (action a_3), thereby directly determining the ability of the entire reinforcement learning model to achieve the objective of anomaly detection. The purpose of action a_2 is to assess the attributes of the adaptive windows obtained under action a_1 and determine whether they contain abnormal signal segments. The same applies to action a_3. Additionally, action a_4 serves as a reliability estimation for action a_2 and aims to assist in the attribute judgment of the adaptive windows. The same applies to action a_5.

Design of state transition probability function P: Since the state space S is designed as a fixed-length set of eddy current amplitude data in a single channel, the state transition process is as follows after executing the corresponding action set: 1) The adaptive window action a_1 promotes the current state change, and the state slides forward in the corresponding channel of the original pipeline amplitude data. 2) The signal classification actions a_2 to

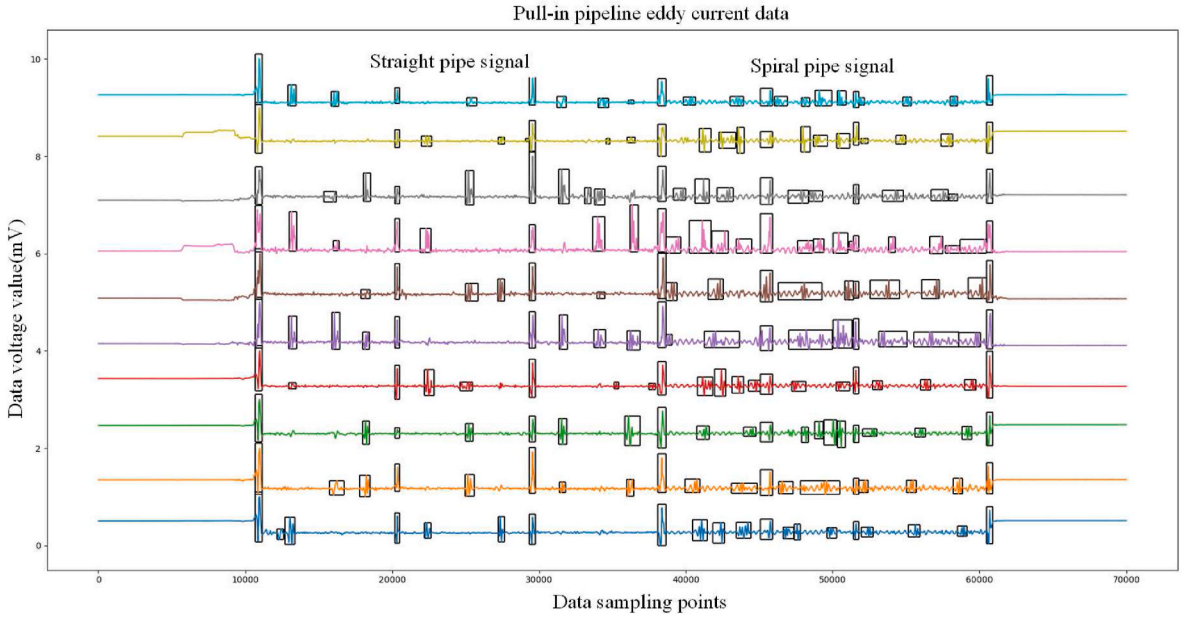
a_5 assess the abnormal attributes and types of the current adaptive window signal, but they do not cause a change in the current state. Therefore, under the joint effect of the above actions set, the state transitions deterministically and the corresponding P should be set to 1.

Design of discount factor γ: In our case, γ is chosen as 0.1 which allows the current rewards more heavily weighted while future long-term rewards are given less importance. This consideration is motivated by the following ground that under the aforementioned parameter design, the adaptive window action of the current state only affects the transition of the current state. It is therefore adjusted by the adaptive window action of the next state, without influencing the transition of the subsequent state or future long-term states. Hence, there is reduced need to overly consider future long-term rewards.

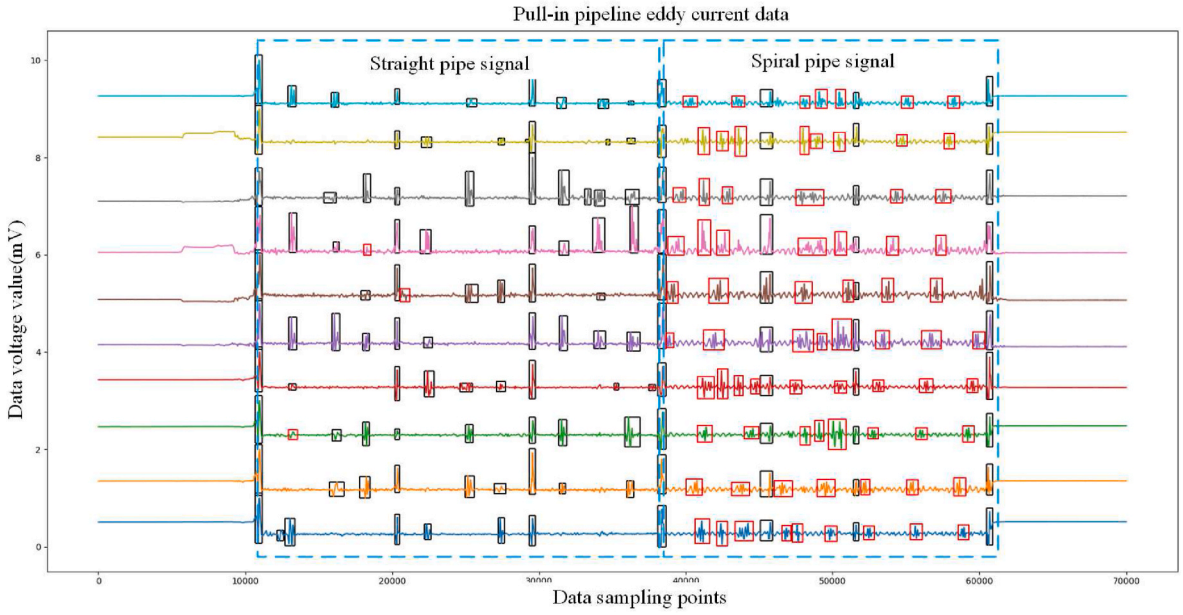
C. Ablation experiments on exploration mechanism

Regarding the adaptive window signal segmentation action, this section conducts a two-step analysis: (1) a visual comparison of pipeline signal window partitioning.

For (1), specific comparisons are shown in Fig. 14, where (a) and (b) illustrate the impact of spatiotemporal exploration mechanisms on adaptive window partitioning for the pull-in pipeline, while (c) and (d) illustrate the impact of spatiotemporal exploration mechanisms on adaptive window partitioning for the field pipeline. It is important to note that the following figure only display the window partitioning corresponding to the abnormal signals, and signals without windows are considered normal signals.

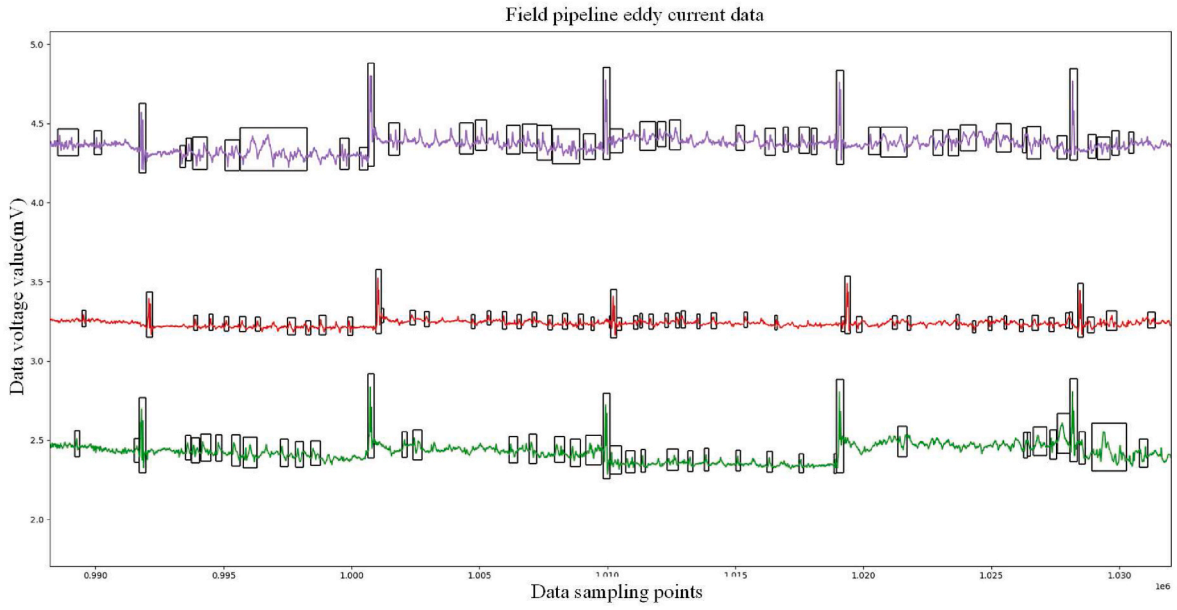


(a)

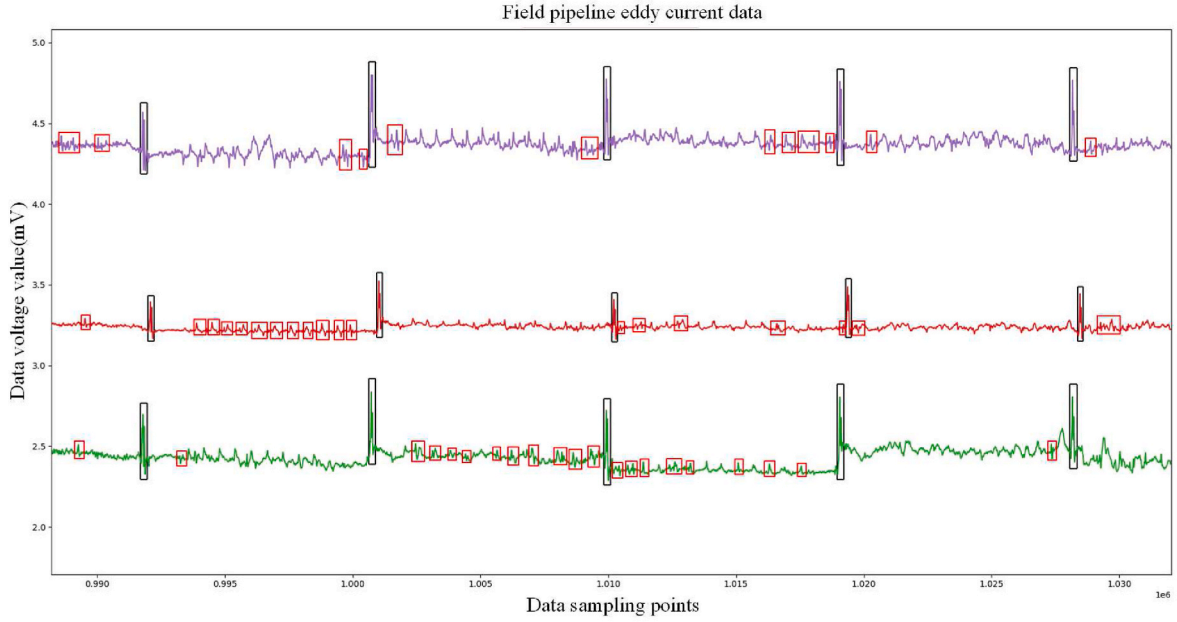


(b)

Fig. 14. Window effects of the adaptive window signal segmentation action (a) Pull-in pipeline without exploration mechanism. (b) Pull-in pipeline with exploration mechanism. (c) Field pipeline without exploration mechanism. (d) Field pipeline with exploration mechanism..



(c)



(d)

Fig. 14. (*continued*).

From Fig. 14(a) and (b), as for the pull-in pipeline, the unoptimized framework shows good window partitioning results for the straight pipe signals (indicated by the blue box on the left side in the figure). The introduction of the spatiotemporal exploration mechanism only slightly fine-tunes the appropriate range of some windows, with no significant overall improvement. However, for the spiral pipe signals (indicated by the blue box on the right side in the figure), the unoptimized framework performs poorly, with some individual anomaly being partitioned into multiple windows or multiple anomalies being assigned to the same window. The introduction of the spatiotemporal exploration mechanism greatly improves these issues, leading to a noticeable overall enhancement in the results.

From Fig. 14(c) and (d) it is evident that the unoptimized framework performs poorly in partitioning abnormal signals for the field pipeline, as it is susceptible to various types of noise interference, leading to erroneous window partitioning, inappropriate window sizes, and unreasonable window distribution. The introduction of the spatiotemporal exploration mechanism enables the optimized framework to accurately recognize complex signals in the field pipeline. The majority of abnormal signals can be accurately and completely partitioned into individual adaptive windows, while the number of incorrectly partitioned windows significantly decreases. The issues observed before optimization are significantly improved, and the effect of adaptive window partitioning is noticeably enhanced.

References

[1] Gupta S, Kumar R, Lu K, et al. Local search methods for k-means with outliers. Proc. VLDB Endowment 2017;10(7):757–68.

[2] Liao L, Luo B. Entropy isolation forest based on dimension entropy for anomaly detection[C]//International Symposium on Intelligence Computation and Applications. Singapore: Springer; 2018. p. 365–76.

[3] Wang J, Miao F, You L, et al. A deep autoencoder based outlier detection for time series[C]//2020 IEEE international conference on power, intelligent computing and systems (ICPICS). IEEE; 2020. p. 216–8.

[4] Bu J, Liu Y, Zhang S, et al. Rapid deployment of anomaly detection models for large number of emerging kpi streams[C]//2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC). IEEE; 2018. p. 1–8.

[5] Kha NH, Anh DT. From cluster-based outlier detection to time series discord discovery[M]. In: Trends and applications in knowledge discovery and data mining. Cham: Springer; 2015. p. 16–28.

[6] Ren H, Xu B, Wang Y, et al. Time-series anomaly detection service at microsoft[C]. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining; 2019. p. 3009–17.

[7] Xu H, Chen W, Zhao N, et al. Unsupervised anomaly detection via variational auto-encoder for seasonal kpis in web applications[C]. In: Proceedings of the 2018 world wide web conference; 2018. p. 187–96.

[8] Zhao N, Zhu J, Wang Y, et al. Automatic and generic periodicity adaptation for kpi anomaly detection. IEEE Transact. Network Serv. Manag. 2019;16(3):1170–83.

[9] Cheng Z, Zou C, Dong J. Outlier detection using isolation forest and local outlier factor[C]//Proceedings of the conference on research in adaptive and convergent systems. 2019. p. 161–8.

[10] Yu Y, Zhu Y, Li S, et al. Time series outlier detection based on sliding window prediction. Math Probl Eng 2014;2014.

[11] Malhotra P, Vig L, Shroff G, et al. Long short term memory networks for anomaly detection in time series[C]//Proceedings, vol. 89. Presses universitaires de Louvain; 2015. p. 89–94.

[12] Ferdousi Z, Maeda A. Unsupervised outlier detection in time series data[C]//22nd international conference on data engineering workshops (ICDEW'06). IEEE; 2006. p. x121. -x121.

[13] Liu Y, Li Z, Zhou C, et al. Generative adversarial active learning for unsupervised outlier detection. IEEE Trans Knowl Data Eng 2019;32(8):1517–28.

[14] Takeuchi J, Yamanishi K. A unifying framework for detecting outliers and change points from time series. IEEE Trans Knowl Data Eng 2006;18(4):482–92.

[15] Teng X, Yan M, Ertugrul AM, et al. Deep into hypersphere: robust and unsupervised anomaly discovery in dynamic networks[C]. In: Proceedings of the twenty-seventh international joint conference on artificial intelligence; 2018.

[16] Liu Z, Qin T, Guan X, et al. An integrated method for anomaly detection from massive system logs. IEEE Access 2018;6:30602–11.

[17] Li C, Chen Z, Liu J, et al. Power load forecasting based on the combined model of LSTM and XGBoost[C]. In: Proceedings of the 2019 the international conference on pattern recognition and artificial intelligence; 2019. p. 46–51.

[18] Yamak PT, Yujian L, Gadosey PK. A comparison between arima, lstm, and gru for time series forecasting[C]. In: Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence; 2019. p. 49–55.

[19] Li J, Yang S. Time series prediction based on decomposition and synthesis[C]. In: Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence; 2019. p. 171–7.

[20] Siami-Namini S, Tavakoli N, Namin AS. The performance of LSTM and BiLSTM in forecasting time series[C]. In: 2019 IEEE international conference on big data (big data). IEEE; 2019. p. 3285–92.

[21] Greff K, Srivastava RK, Koutník J, et al. LSTM: a search space odyssey. IEEE Transact Neural Networks Learn Syst 2016;28(10):2222–32.

[22] Yang Y, Gao B, Liu D, et al. Electromagnetic pigging system based on sandwich differential planar coil. IEEE Sensor J 2022;22(19):18903–13.

[23] Roulston MS, Smith LA. Evaluating probabilistic forecasts using information theory. Mon Weather Rev 2007;130(6):1653–60.

[24] Ren Liqiang, Jia Shuyi, Wang Haipeng, Wang Ziling. A review of research on time series classification based on deep learning[J]. J Electron Inf Technol. doi: 10.11999/JEIT231222.

[25] Karim F, Majumdar S, Darabi H, et al. Multivariate LSTM-FCNs for time series classification. Neural Network 2019;116:237–45.

[26] Cui Z, Chen W, Chen Y. Multi-scale convolutional neural networks for time series classification 2016. arXiv preprint arXiv:1603.06995.