A Novel Deep Offline-to-Online Transfer Learning Framework for Pipeline Leakage Detection with Small Samples

Chuang Wang, Zidong Wang, *Fellow, IEEE*, Weibo Liu, Yuxuan Shen, and Hongli Dong, *Senior Member, IEEE*

***Abstract*—In this paper, a two-stage deep offline-to-online transfer learning framework (DOTLF) is proposed for long-distance pipeline leakage detection (PLD). At the offline training stage, a feature transfer-based long short-term memory network with regularization information (TL-LSTM-Ri) is developed where a maximum mean discrepancy regularization term is employed to extract domain-invariant features and an adjacent-bias-corrected regularization term is introduced to extract early fault features from pipeline samples under different scenarios. At the online detection stage, the trained TL-LSTM-Ri is employed for motion prediction so as to monitor the operating condition of the pipeline in real time. To demonstrate its application potential, the DOTLF is successfully applied to handle the PLD problem on the long-distance oil-gas pipeline data. Experimental results demonstrate the effectiveness of the proposed DOTLF for real-time PLD under real-world scenarios.**

***Index Terms*—Deep transfer learning, dynamic threshold, long short-term memory network, pipeline leakage detection, small samples.**

I. Introduction

A pipeline is an energy-efficient carrier for transporting fluids (e.g., oil, water, and natural gas) with high commercial value, which plays a significant role in modern petroleum industry [45]. Due to the corrosive and flammable nature of petroleum, the failure of pipeline (such as leakage and blockage) would occur, which could lead to economic loss and serious safety hazards (e.g., environmental pollution and loss of human life) [32], [47]. As such, it is of vital importance to carry out fault detection to ensure the safe operation of the pipeline [3], [7], [14], [27], [37], [38], [49], [51], [52], [59].

The fault detection of pipeline and pipe network has attracted an ever-increasing interest from both academia and industry [32], [45]. A great many data analysis methods have been developed for fault detection [1]. Among them, modern deep learning (DL) techniques have been successfully exploited in pipeline leakage detection (PLD) thanks to the powerful feature extraction ability [20], [26], [30], [35], [39], [54], [58]. Although some DL-based PLD methods exhibit promising performance, there are two major requirements to train an effective DL model: 1) a large number of samples with high-quality annotations are needed for model training; and 2) the training and testing samples should obey the same data distribution. Unfortunately, it is nearly impossible to meet the aforementioned requirements in real-world PLD scenarios since the pipeline at abnormal state will be maintained immediately, and it is expensive and time-consuming to manually annotate the samples [13].

Transfer learning (TL) provides a useful framework for tackling the insufficient training data problem, which aims to transfer the knowledge from the source domain to the target domain [36]. Leveraging the power of TL, a great number of industrial fault detection methods have been put forward in the past few years [4], [9], [11], [16], [17], [19], [22], [23], [40], [42]. Owing to the lack of large-scale annotated real-world pipeline datasets, a commonly used way is to simulate the pipeline operation in the laboratory to obtain sufficient source domain data [56]. In fact, there might be a large discrepancy between the distribution of the pipeline data collected under different operating conditions and working environments, which would cause the domain shift problem.

Fortunately, a variety of domain adaption-based TL methods have been proposed to overcome the domain shift problem for fault detection [9], [16], [23], [42], [56]. For instance, a deep convolutional transfer learning network has been developed in [9] to deal with the problem of bearing data without label by not only maximizing the domain recognition errors but also minimizing the probability distribution distance. Recently, a multi-layer domain adaptation method has been proposed in [56] to reduce the distribution discrepancy and narrow the inter-class distance of the transferable features for tackling the problem of bearing fault diagnosis. More recently, a residual joint adaptation adversarial network has been presented in [16] for intelligent fault detection, where the joint distribution discrepancy has been considered in transferring the detection knowledge. In this context, the domain adaption-based TL method is chosen to handle the PLD problem with small samples in this paper.

It should be mentioned that most existing DL-based fault

This work was supported in part by the National Natural Science Foundation of China under Grants 61933007, U21A2019, and 62103095, the Hainan Province Science and Technology Special Fund of China under Grant ZDYF2022SHFZ105, the Natural Science Foundation of Heilongjiang Province of China under Grant LH2021F005, and the Alexander von Humboldt Foundation of Germany. (*Corresponding author: Hongli Dong.*)

Chuang Wang and Yuxuan Shen are with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing 163318, China, and also with the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Northeast Petroleum University, Daqing 163318, China (e-mails: wangchuang64@126.com, shenyuxuan5973@163.com).

Zidong Wang and Weibo Liu are with the Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, U.K. (e-mails: Zidong.Wang@brunel.ac.uk, Weibo.Liu2@brunel.ac.uk).

Hongli Dong is with Sanya Offshore Oil & Gas Research Institute, Northeast Petroleum University, Sanya 572025, China, and also with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing 163318, China (e-mail: shiningdhl@vip.126.com).

detection methods mainly exploit fault features (which are extracted from historical data) to classify the input data into normal or fault states. Nevertheless, these methods may not trigger alarms in time when encountering unseen faults. In this case, the aforementioned DL-based fault detection methods cannot meet the requirement of real-world PLD under complex scenarios. To tackle the aforementioned PLD problem, we aim to develop a novel DL-based fault detection method to predict the pipeline pressure, where an alarm would trigger when the real-time sensor data exceeds the dynamic threshold of the predicted value.

According to the above discussions, we are motivated to solve the real-time PLD problem with small samples. The following key challenges need to be addressed: 1) how to build a reliable real-time PLD method; 2) how to train an effective detection model with insufficient data; and 3) how to guarantee the detection accuracy of the designed PLD method. Pertaining to the aforementioned challenges, we endeavor to propose a novel two-stage deep offline-to-online transfer learning framework (DOTLF) for real-world PLD in this paper. In the proposed DOTLF, a feature transfer-based long short-term memory network with regularization information (TL-LSTM-Ri) is proposed to extract domain-invariant features so as to improve the prediction accuracy and solve the small sample problem at the offline training stage. At the online detection stage, the developed TL-LSTM-Ri is employed for motion prediction to monitor the working condition of the pipeline in real time. The main contributions of this paper are summarized as follows.

1) A novel DOTLF is proposed for real-time PLD, which can not only tackle the small sample problem but also detect unseen faults with high accuracy.
2) A novel feature transfer-based LSTM-Ri algorithm is developed where a maximum mean discrepancy-based regularization term is employed to alleviate the domain shift problem and an adjacent bias-corrected regularization term is designed to extract normal condition features and early fault features.
3) The proposed DOTLF is successfully exploited in analyzing real-world oil-gas pipeline data and detecting pipeline leakage in real time under complex industrial scenarios. Experimental results demonstrate that the proposed DOTLF exhibits competitive or even superior performance than some oil-gas PLD methods in terms of detection accuracy.

The rest of this paper is organized as follows. The background of the LSTM and the TL are introduced in Section II. The details of the developed DOTLF and the TL-LSTM-Ri are introduced in Section III. Experiment setting and results are presented in Section IV. Finally, the conclusions are drawn, and the future work is pointed out in Section V.

II. Preliminaries

A. Long Short-Term Memory Network

The recurrent neural network (RNN) has become a popular algorithm for dealing with time-dependent problems in the past few decades, see [15], [25], [43] and the references therein. The structure of a standard RNN is displayed in Fig. 1. As shown in Fig. 1, the output of the hidden layer is determined by the current input and the output of the previous hidden layers.

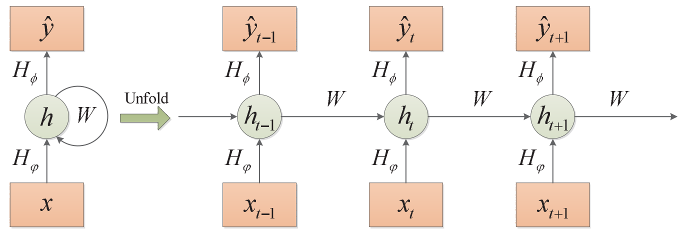


Fig. 1: The structure of the standard RNN.

To handle the vanishing gradient and exploding gradient problems, various DL methods have been recently proposed such as the LSTM [12], the gated recurrent unit (GRU) [5], the temporal convolutional network (TCN) [2]. Compared to the GRU and the TCN, the LSTM exhibits competitive or even superior performance thanks to its network architecture. The structure of a single storage unit in the standard LSTM is shown in Fig. 2.

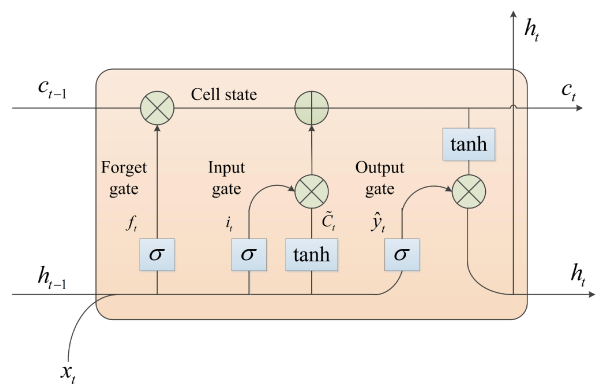


Fig. 2: The structure of a single storage unit of the LSTM.

The key to the LSTM structure is to design the cell state as shown in Fig. 2. The cell can be regarded as a conveyor belt, which directly transmits the information obtained from the previous layers through the whole chain to the current layer with linear interactions. Generally, three gates are designed to store/abandon information to the cell state in the LSTM. It can be seen in Fig. 2 that a single storage unit of the standard LSTM contains a forget gate, an input gate, and an output gate. The forget gate aims to determine how much information from the previous layers should be discarded at the current layer by using the sigmoid function. The forget gate f_t is described as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \tag{1}$$

where $\sigma(\cdot)$ is the standard sigmoid function; t is the current iteration; W_f and b_f represent the weight and bias of the forget gate, respectively; x_t is the input data; and h_{t-1} denotes the output of hidden layer at the $(t-1)$th iteration.

The input gate i_t is employed to determine how much information from the new input should be preserved in the current cell state, which can be formulated as (2). Meanwhile, the new state information $\tilde{C}_t$ is obtained by (3). As such, the current cell state C_t can be updated by (4):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \tag{2}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \tag{4}$$

where $*$ represents the point-wise multiplication; W_i is the weight of the input gate; b_i is the bias of the input gate; W_c is the weight of the new state information; and b_c represents the bias of the new state information.

The purpose of the output gate is to decide how much information should be propagated to the next layer. In general, the output gate $\hat{y}_t$ is updated as follows:

$$\hat{y}_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \tag{5}$$

$$h_t = \hat{y}_t * \tanh(C_t), \tag{6}$$

where C_t denotes the cell state at the tth iteration; W_o and b_o are the weight and the bias of the output gate, respectively; h_t denotes the output of the LSTM cell at the tth iteration.

B. Transfer Learning

Transfer learning (TL) provides a framework where a system could apply knowledge/skills obtained from previous tasks to new tasks, which can be used to handle few-shot or even zero-shot problems [34], [53]. In TL, two datasets are drawn from the source domain and target domain, respectively. The source data aims to provide the detection knowledge to train the models. The target data makes full use of the transferred knowledge with the purpose of solving classification or regression problems with a limited number of samples. In fact, the TL methods may not perform well without taking the correlation between source data and target data into account, which leads to the phenomenon of negative transfer. As such, it is necessary to fully consider the correlation between source data and target data when using the TL methods.

III. A Novel Offline-to-Online Transfer Learning Framework

This paper aims to develop an intelligent PLD method by transferring the acquired fault diagnosis knowledge from the pipeline samples under laboratory conditions (PLCs) to the pipeline samples under real-case conditions (PRCs). A two-stage offline-to-online transfer learning framework is designed for PLD. In the proposed framework, the leak detection process is divided into the offline training stage and the online detection stage.

A. Problem Formulation

In this paper, we aim to perform a two-stage offline-to-online PLD based on the transfer learning framework, where the PLD knowledge acquired from the PLCs is employed for data analysis on the PRCs. Let $\mathcal{D}_s$ and $\mathcal{D}_t$ denote the source domain and the target domain, respectively.

Suppose that we have the sample space $\mathcal{X}_s \in \mathcal{D}_s$ and $\mathcal{X}_t \in \mathcal{D}_t$, the data sampled from the source and target domains are denoted by $x_s \in \mathcal{X}_s$ and $x_t \in \mathcal{X}_t$, respectively. The source data for offline training and the target data for online detection obey the marginal probability distribution $P(x_s)$ and $P(x_t)$, respectively.

The source domain $\mathcal{D}_s$ consists of the sample space $\mathcal{X}_s$ and the distribution $P(x_s)$ of PLCs, which can be represented by $\mathcal{D}_s = \{\mathcal{X}_s, P(x_s)\}$. The target domain $\mathcal{D}_t$ consists of the sample space $\mathcal{X}_t$ and the distribution $P(x_t)$ of PRCs, which can be denoted by $\mathcal{D}_t = \{\mathcal{X}_t, P(x_t)\}$.

Due to the lack of annotated target domain data, it is necessary to apply the detection knowledge acquired from the source domain to assist the data analysis on the target domain data. In this context, the source domain $\mathcal{D}_s$ should provide sufficient detection knowledge for PLD of the target domain $\mathcal{D}_t$. It should be noted that both $\mathcal{D}_s$ and $\mathcal{D}_t$ should include the pipeline operation information under normal conditions and leakage conditions.

Remark 1: Different from traditional fault detection problems, the online pipeline leakage detection can be treated as an outlier detection problem [41], [61]. Note that most online detection models tend to fail in identifying incipient faults due to insufficient real-world pipeline data. Therefore, this paper aims to put forward a novel two-stage TL framework for PLD, where sufficient detection knowledge is acquired at the offline training stage by analyzing PLCs for online detection. The proposed TL-based PLD framework could: 1) effectively capture the domain-invariant features from the source data and target data, which significantly improves the generalization ability of the detection model; and 2) extract the features under the normal condition and the fault-working condition, which contributes to the fault detection at early stage.

B. Overview of the DOTLF

In this paper, the DOTLF is put forward for real-time PLD. The flowchart of the proposed DOTLF is shown in Fig. 3. At the offline training stage, the TL-LSTM-Ri is presented where a modified loss function is developed to tackle the small samples problem and the high false-alarm rate problem. To be specific, the mean square error (MSE), the maximum mean discrepancy (MMD) regularizer, and an adjacent bias-corrected (AB) regularizer are employed to construct the modified loss function. Owing to the developed loss function, the proposed TL-LSTM-Ri is capable of 1) adaptively extracting the common feature representation of the PRCs and the PLCs; and 2) effectively distinguishing the features extracted from the normal state and the incipient fault state. At the online detection stage, the trained TL-LSTM-Ri is employed for motion prediction to monitor the working condition of the pipeline in real time, which can adaptively tune the threshold under different conditions.

C. Offline Training Stage

In this subsection, a novel TL-LSTM-Ri is put forward where a novel loss function is developed by introducing two regularization terms to the traditional loss function of the

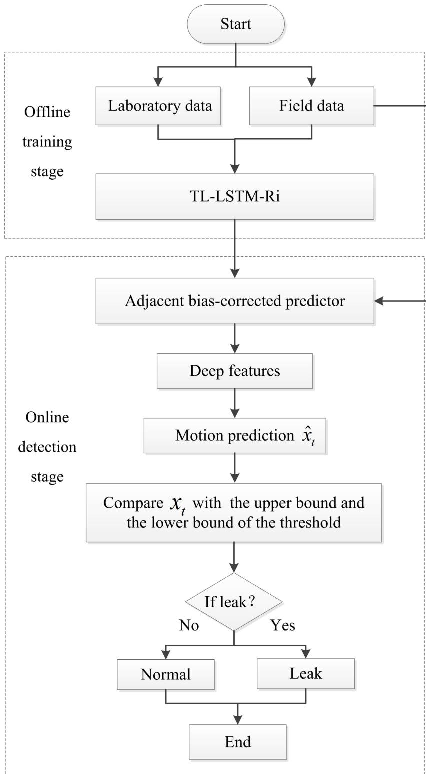


Fig. 3: The flowchart of the DOTLF algorithm.

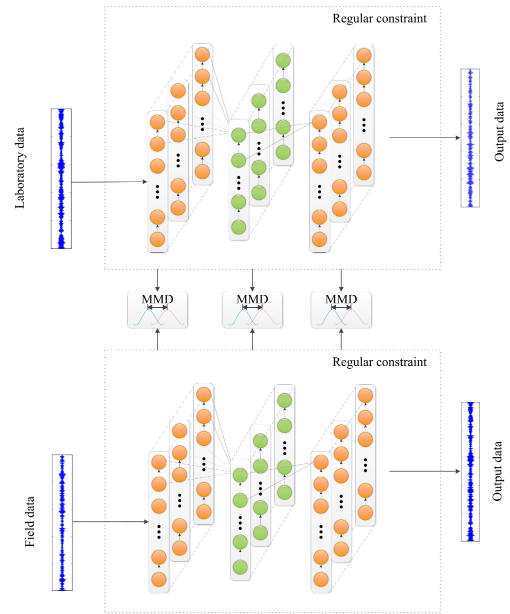


Fig. 4: The structure of the TL-LSTM-Ri.

LSTM. To be specific, the introduction of the MMD aims to reduce the distribution discrepancy between the source domain and the target domain. The utilization of the AB regularization term focuses on improving the discrimination between the incipient fault features and the normal features. The structure of the TL-LSTM-Ri is displayed in Fig. 4.

1) Motivation: The RNN algorithm and its variants (especially the LSTM) have achieved a great success in dealing with time-series problems. Note that the traditional LSTM is point-to-point prediction. Nevertheless, the prediction accuracy tends to be negatively impacted by some outliers (anomalies but not faults). Under this circumstance, a seemingly natural idea is to develop an *adjacent bias-corrected* predictor, which could reduce prediction bias by making the predicted data more similar to its neighboring regions. In this paper, a novel TL-LSTM-Ri is proposed for high-accuracy pipeline leakage prediction. The proposed TL-LSTM-Ri could analyze the feature information about the trend of the pipeline pressure by using the designed AB regularization term, which benefits the fault detection on early minor faults.

2) Loss Function: The proposed loss function includes three parts which are the loss function of the traditional LSTM, the MMD, and the AB regularization term. The loss function of the traditional LSTM (i.e., MSE) is given as follows:

$$L_1(x_s;\theta) = \frac{1}{n_s}\sum_{i=1}^{n_s}\left(\hat{x}_s(i) - x_s(i)\right)^2, \tag{7}$$

where $x_s(i)$ represents the ith source data; $\hat{x}_s(i)$ denotes the ith predicted data; θ represents the parameters in the LSTM; and n_s represents the total number of the source data.

With the purpose of eliminating the prediction bias, a novel regularization term is designed. The loss of the AB term is given as follows:

$$L_2(x_s;\theta) = \ln\left(\frac{1}{n_s}\sum_{i=1}^{n_s}\left(\hat{x}_s(i) - \frac{1}{n}\sum_{\tau=1}^{n} x_s(i-\tau)\right)^2\right), \tag{8}$$

where τ represents a known constant; n denotes the upper bound of the known constant; n_s represents the total number of the source domain samples. Note that the loss of the AB regularization term is calculated according to (8) when τ is smaller than i, and otherwise we set $\tau = 0$.

In order to capture the domain-invariant features, the MMD is employed in this paper [6], [33]. Unlike the Kullback-Leibler (KL) divergence, the MMD is suitable for estimating the non-parametric distance among different distributions by means of the reproducing kernel Hilbert space (RKHS) [31]. The MMD distance is given as follows:

$$\mathrm{MMD}(x_s, x_t) = \left\| \frac{1}{n_s}\sum_{i=1}^{n_s}\varphi(x_s(i)) - \frac{1}{n_t}\sum_{j=1}^{n_t}\varphi(x_t(j)) \right\|_{\mathcal{H}}, \tag{9}$$

where $\mathcal{H}$ denotes the RKHS; $\varphi(\cdot)$ is the nonlinear mapping from the original feature space to the RKHS; x_s and x_t

represent the input samples obtained from the source data and target data, respectively; n_s and n_t stand for the total number of samples in the source data and the target data, respectively. In this paper, an MMD-based regularizer is designed to reduce the distribution discrepancy between source data and target data. The loss function of the designed MMD regularizer is given by:

$$L_3(x_s, x_t; \theta) = \frac{\mathrm{MMD}(x_s, x_t)}{1 + \mathrm{MMD}(x_s, x_t)}. \tag{10}$$

In this paper, the proposed loss function of the TL-LSTM-Ri is given by:

$$L(x_s, x_t; \theta) = (L_1(x_s; \theta) + \gamma L_2(x_s; \theta)) L_3(x_s, x_t; \theta), \tag{11}$$

where γ is a constant value to balance the MSE loss function and the AB loss function.

Remark 2: The main novelty of the proposed TL-LSTM-Ri lies in the design of a novel loss function for motion prediction based on the transferred knowledge. Compared with the classic MMD constraint, a fractional form of the MMD is designed as a regularization term. The designed MMD-based regularization term approaches to 1 with the increase of the MMD, and the designed term trends to reach 0 as MMD decreases. Such a fractional MMD regularization term acts like a control parameter, which could influence the knowledge transfer between the source domain and the target domain. The designed AB regularization term acts as a "smoothing" factor, which would result in the moving prediction rather than the point-to-point prediction for real-time fault detection.

In this paper, the Adam gradient descent algorithm is employed to minimize (11). Letting $f(x_s, x_t; \theta) = \mathrm{MMD}(x_s, x_t)$, $g(x_s; \theta) = L_1(x_s; \theta) + \gamma L_2(x_s; \theta)$, one has:

$$L(x_s, x_t; \theta) = \frac{f(x_s, x_t; \theta)}{1 + f(x_s, x_t; \theta)} g(x_s; \theta). \tag{12}$$

The gradient is computed by:

$$\begin{aligned} \frac{\partial L(x_s, x_t; \theta)}{\partial \theta} &= \frac{g(x_s; \theta)}{(f(x_s, x_t; \theta) + 1)^2} \frac{\partial f(x_s, x_t; \theta)}{\partial \theta} \\ &\quad + \frac{f(x_s, x_t; \theta)}{1 + f(x_s, x_t; \theta)} \frac{\partial g(x_s; \theta)}{\partial \theta}. \end{aligned} \tag{13}$$

During the training process, the weighting parameters θ of the proposed TL-LSTM-Ri are updated as follows:

$$\theta \leftarrow \theta - \alpha \frac{\partial L(x_s, x_t; \theta)}{\partial \theta}, \tag{14}$$

where α is the learning rate.

D. Online Pipeline Leakage Detection Stage

The process of online detection is illustrated in Fig. 5. First of all, the real PRCs (depicted in black) for the first $T - 1$ time points are fed into the trained TL-LSTM-Ri model to predict the PRCs (depicted in red) at the moment T. Then, the predicted PRCs at the moment T are compared with the corresponding real PRCs to judge the pipeline condition according to the dynamic threshold. Note that if the sample points at moment T are under fault condition, then the predicted PRCs at moment T will replace the real PRCs as the input of the detection model to predict the PRCs at moment $T + 1$. Conversely, the real PRCs at moment T will continue to be used to predict the PRCs at moment $T+1$, and the cycle repeats until the end. In summary, the prediction mode of the DOTLF can be discussed in two cases: 1) if the current sample is normal, the DOTLF can be considered as a one-step-ahead prediction; and 2) if the current sample is faulty, the DOTLF can be regarded as a combination of one-step-ahead prediction and multi-step-ahead prediction.

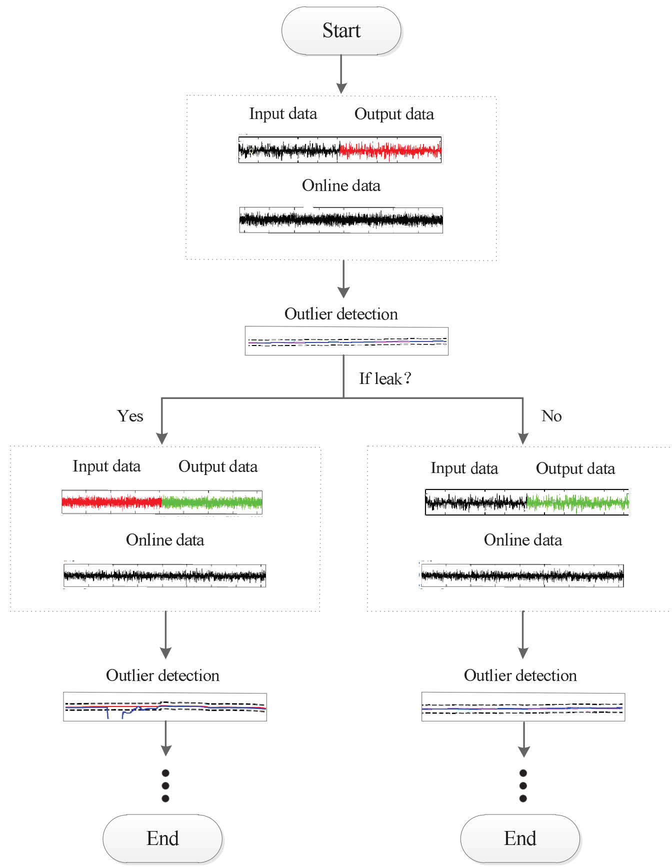


Fig. 5: Process of online detection.

IV. Experiments and Results

In this paper, experiments are carried out according to the following four steps: 1) data collection and pre-processing; 2) model pre-training; 3) diagnostic knowledge transfer; and 4) real-time fault diagnosis. In the experiment, a large number of PLCs collected from the laboratory are chosen as the source data and a limited number of PRCs are employed as the target data. In the following subsections, we discuss the details about the data pre-processing, the offline training process as well as the online detection process.

A. Data Preprocessing

It is known that the TL technique could reduce the distribution discrepancy between the source and target domains. In this paper, the source data and target data are chosen from the PLCs and PRCs, respectively. The experimental platform for simulating the real-world pipeline data is depicted in Fig. 6. A simulation sample is shown in Fig. 7. The parameters of the experimental system are illustrated in the following: the length of the pipeline is 180m, there is a leakage point every 10m, the pressure is 0.5MPa, and the flow rate is 60m^3/h. The PRCs are collected from the oil pipeline network system. A real-world sample is the negative pressure wave signal which is shown in Fig. 8. It can be observed in Figs. 7-8 that the amplitude difference between the two signals reaches 6 orders of magnitude, which would affect the knowledge transfer performance. Additionally, owing to the influence of the sensors and the outside environment (e.g., moving vehicles and site construction), the collected pipeline pressure signal is often corrupted or even drowned by various noises, which gives rise to omission and misreporting of possible accidents. To overcome the aforementioned weaknesses, the normalization and denoising strategies proposed in [46] are employed to alleviate the noise interference.

Fig. 6: The experimental platform of the pipeline leakage detection.

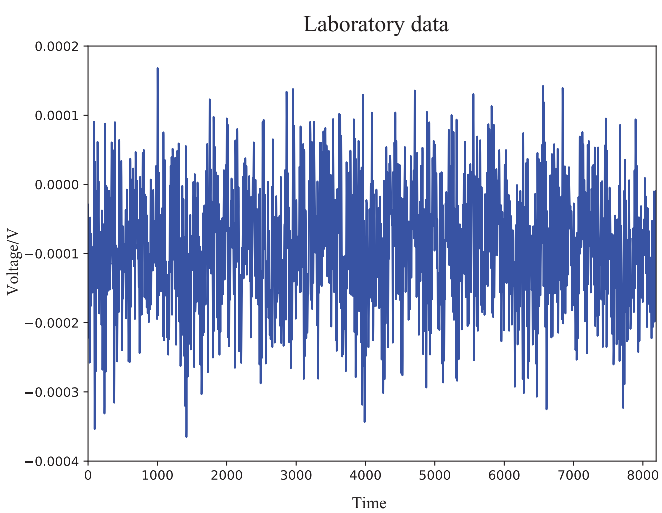


Fig. 7: Laboratory data.

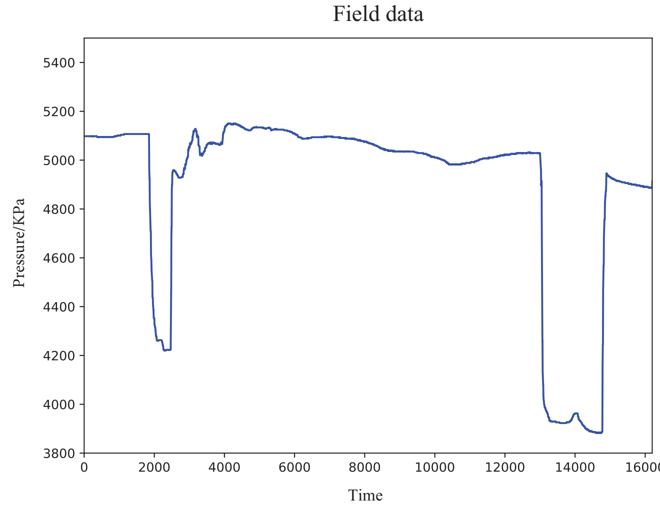


Fig. 8: Field data.

B. Experiment Results of the Offline Training Stage

In this section, the performance of the proposed TL-LSTM-Ri is evaluated by comparing with the standard LSTM, the TL-LSTM, and the LSTM-Ri. For implementation of the proposed TL-LSTM-Ri, a four-layer LSTM is employed to extract features from the pipeline data followed by a three-layer fully connected network for domain adaptation and prediction. For fairness of comparison, the identical network structure and parameters are utilized in all the algorithms. The batch size is set to be 200. The learning rate α of the optimizer is set to be $1e-4$. The step size is set to be 12. The iteration number is set to be 2500, and γ is set to be $1e-3$.

The experimental results of the LSTM, the LSTM-Ri, the TL-LSTM, and the proposed TL-LSTM-Ri at the offline training stage are shown in Fig. 9. Intuitively, as shown in Fig. 9(a) and Fig. 9(c), there is a large gap between the predicted data and the real data. In contrast, it can be seen clearly that the predicted data completely tracks the real-time data in Fig. 9(b) and Fig. 9(d). Notably, the real data are described by the blue curves, and the predicted data are denoted by the red curves. The performance of the utilized algorithms is evaluated according to the difference between the predicted and real data. It can be found that the proposed TL-LSTM-Ri achieves the smallest error among the four algorithms, which indicates the superiority of the proposed TL-LSTM-Ri. By introducing the MMD, the proposed TL-LSTM-Ri exhibits better feature extraction performance than other utilized algorithms.

The loss of the proposed TL-LSTM-Ri is described in Fig. 10, where the vertical coordinate indicates the loss value, and the horizontal coordinate represents the number of iterations. As shown in Fig. 10, the loss of the proposed TL-LSTM-Ri decreases gradually and converges very fast, which indicates the fast convergence and robustness of the TL-LSTM-Ri at the offline training stage. To summarize, the proposed TL-LSTM-Ri outperforms the compared algorithms in terms of prediction performance.

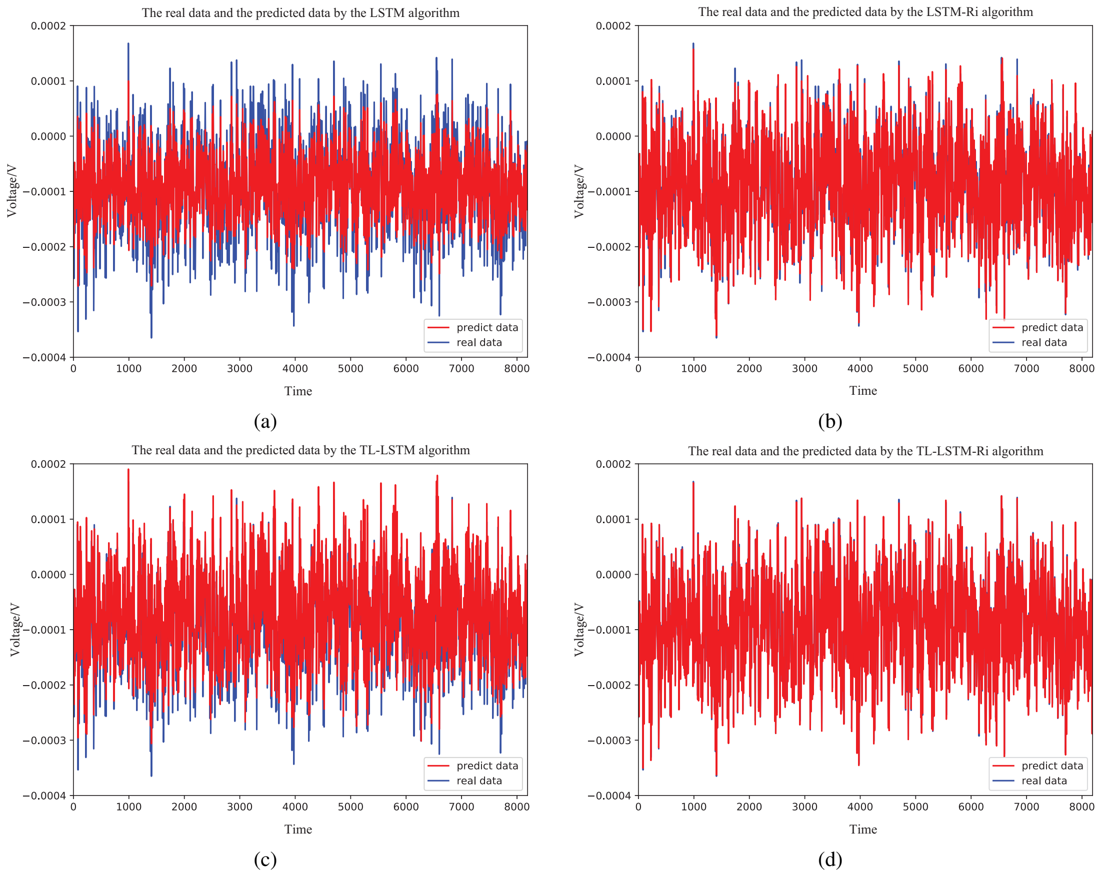


Fig. 9: The real data and the predicted data obtained by using a) LSTM; b) LSTM-Ri; c) TL-LSTM; and d) TL-LSTM-Ri.

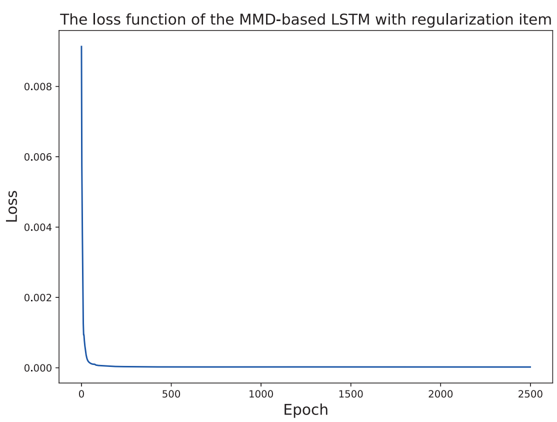


Fig. 10: The loss function of the MMD-based LSTM with regularization terms.

C. Experiment Results of the Online PLD Stage

The performance of the online PLD system is evaluated under both fault and normal conditions. The parameters of the selected algorithms remain the same at the offline PLD stage. In our experiments, the upper bound on the dynamic threshold (UDT) is set as $\mathrm{UDT} = \mathrm{D}_{prediction} \times (1 + 2\%)$, and the lower bound on the dynamic threshold (LDT) is set as $\mathrm{LDT} = \mathrm{D}_{prediction} \times (1 - 2\%)$, where $\mathrm{D}_{prediction}$ is the predicted value obtained by the employed algorithm.

1) Case 1: Online PLD System for Fault Conditions: The four online PLD systems under fault condition have been constructed by using the selected algorithms, which are shown in Fig. 11. The constructed online PLD systems can be roughly divided into the direct transfer learning (DTL) ones and the adaptive transfer learning (ATL) ones. To be specific, the two DTL-based systems are established by using the LSTM and the LSTM-Ri, and the ATL-based systems are constructed based on the TL-LSTM and the TL-LSTM-Ri.

As displayed in Fig. 11, the blue curve represents the field data collected from the oil pipeline network system. The red curve represents the predicted sample, and two black dotted lines are the upper and lower bounds of the dynamic thresholds. The sampling frequency of the oil pipeline network system is 1024Hz. In the experiment, the pipeline leakage occurs at a certain time. The signal size is 16384 and the time window is set to be $T = 16$s.

It can be clearly seen in Fig. 11(a) and Fig. 11(b), the two online PLD systems with the DTL-based LSTM and the DTL-based LSTM-Ri are employed to detect the incipient leakage, respectively. The online detection systems based on the ATL are utilized to identify the leakage as shown in Fig. 11(c) and Fig. 11(d). It can be seen that there is a large error between the predicted data of the DTL and the real data, which indicates high false-alarm rate, see enlarged Figs. 12-13 for details.

Comparing with the DTL-based PLD systems, the ATL-based systems obtain less error between the predicted and real

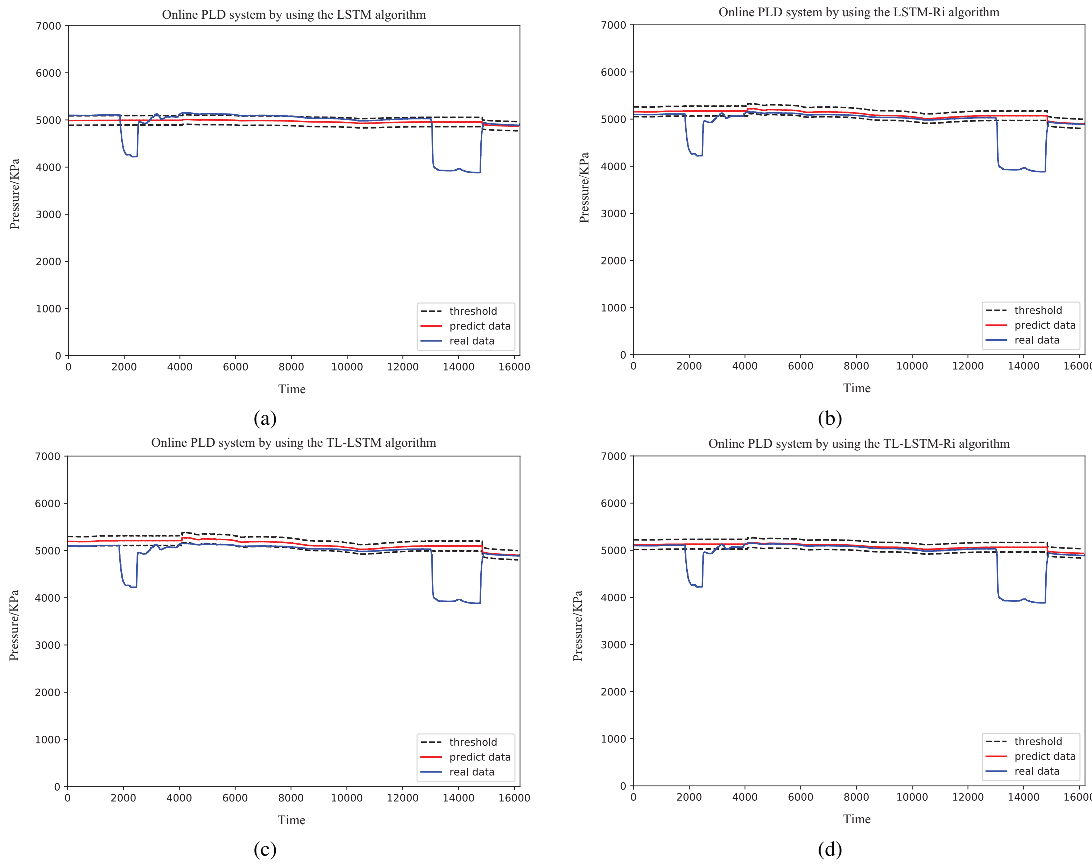


Fig. 11: The online PLD systems by using a) LSTM; b) LSTM-Ri; c) TL-LSTM; and d) TL-LSTM-Ri.

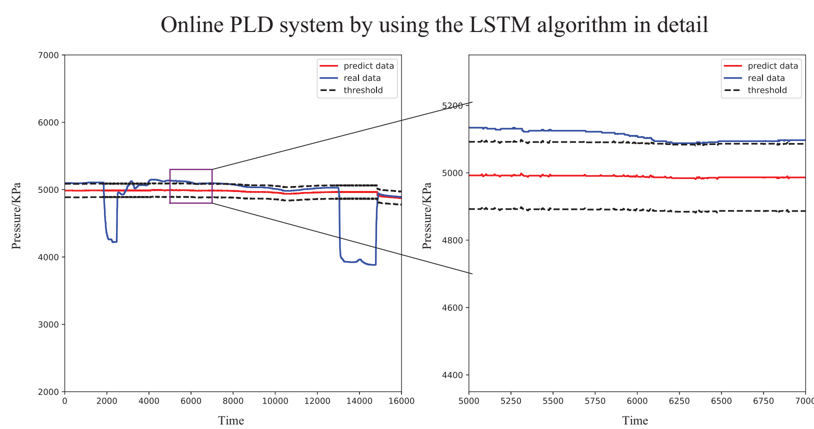


Fig. 12: The online PLD system built by using the LSTM algorithm.

data, which could detect the incipient leakages with less or no false alarm as shown in Figs. 14-15. To summarize, the TL-LSTM-Ri-based PLD system outperforms other designed PLD systems, which could not only detect the leakages but also greatly reduce the false-alarm rate.

2) Case 2: Online PLD System for Normal Conditions: To monitor the pipeline operation under normal conditions, the proposed ATL-based LSTM-Ri is employed to build the online PLD system. As shown in Figs. 16-17, the predicted data could "perfectly" track the real data within the bounds, indicating that the proposed online PLD system is able to alleviate the omission and the misreporting problems.

D. Generalization Ability Evaluation

In previous experiments, we have verified the superior performance of the DOTLF algorithm, which utilizes knowledge of the source domain (laboratory voltage data) to train a detection model for the target domain (field pressure data). In this experiment, we further evaluate the generalization ability of our algorithm in another target domain. Specifically, we test our algorithm with pipeline samples from flow sensors, as shown in Figs. 18-19. First of all, we can observe that

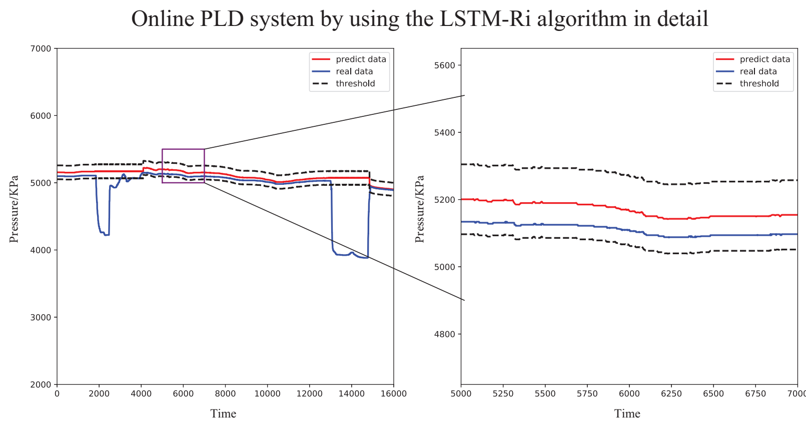


Fig. 13: The online PLD system built by using the LSTM-Ri algorithm.

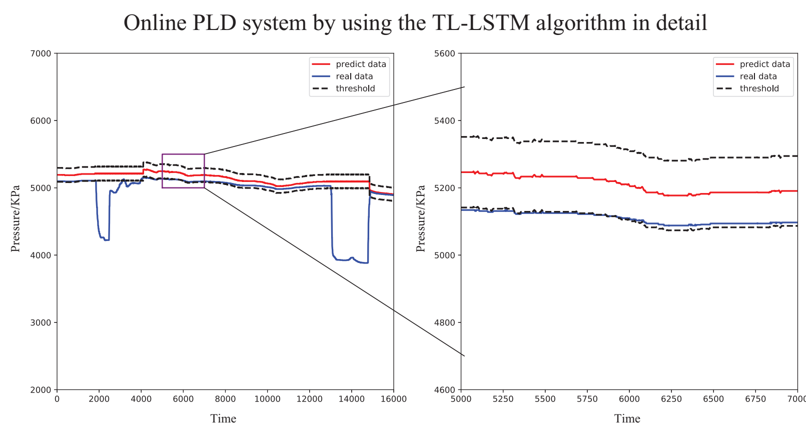


Fig. 14: The online PLD system built by using the TL-LSTM algorithm.

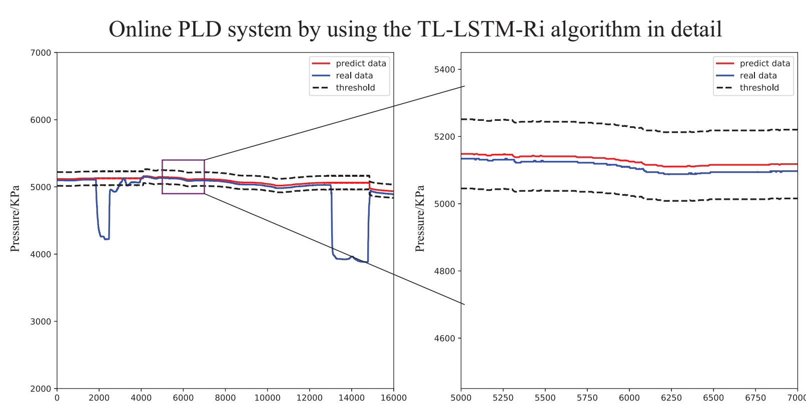


Fig. 15: The online PLD system built by using the TL-LSTM-Ri algorithm.

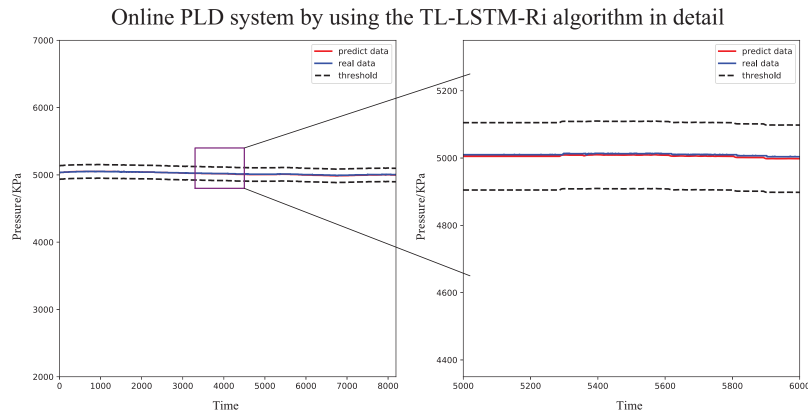


Fig. 17: The TL-LSTM-Ri-based online PLD system under normal conditions: An enlarged case.

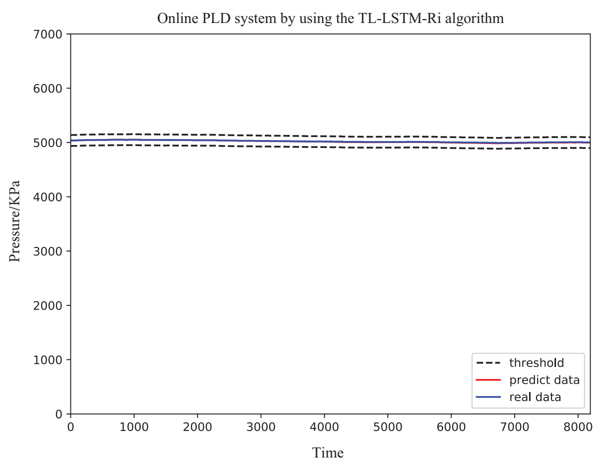


Fig. 16: The TL-LSTM-Ri-based online PLD system under normal conditions.

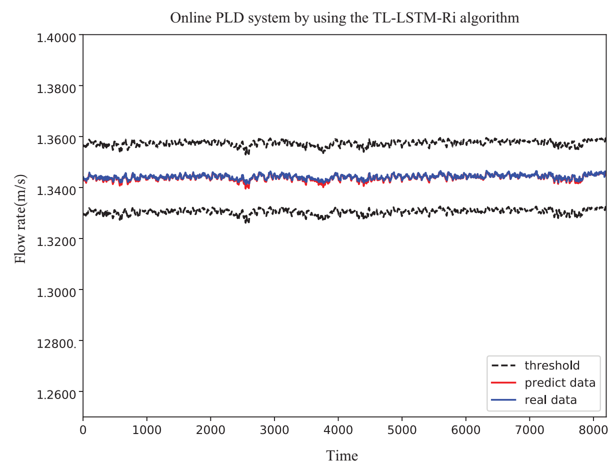


Fig. 18: Online PLD system based on TL-LSTM-Ri algorithm operating under normal condition.

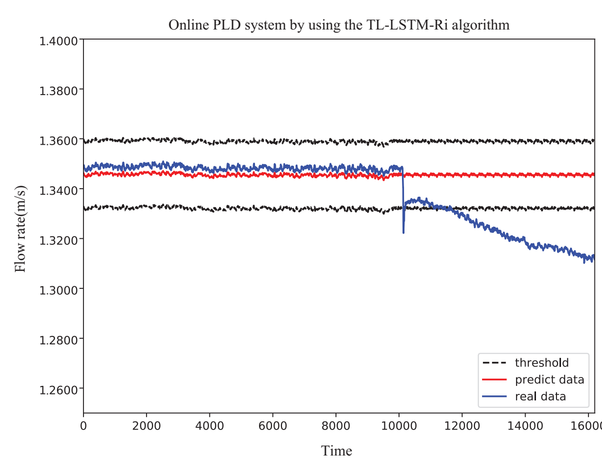


Fig. 19: Online PLD system based on TL-LSTM-Ri algorithm operating under fault condition.

the predicted data precisely tracks the real data under normal condition without any abnormal trends, which verifies that our method can guarantee prediction accuracy, even in the unseen domain. Simultaneously, the DOTLF is capable of accurately detecting incipient leakages under fault condition with few or no false alarms, which demonstrates that our method can achieve expected detection accuracy in the unseen domain. Clearly, our superior generalization performance remains even in the case of more challenging unseen domains.

E. Comparison with the State-of-the-Art

Additional experiments on statistical measurements are implemented in this section to further verify the superiority of our TL-LSTM-Ri algorithm.

1) Baseline Algorithms: To comprehensively evaluate the superiority of the proposed TL-LSTM-Ri algorithm, we compare its performance with that of four state-of-the-art TL methods, including CNN-MMD [24], DANN [8], DeepCORAL [44], and JAN [29]. Specifically, except for the CNN-MMD algorithm that is constructed via convolutional neural networks, all other compared methods share the same network and parameter configuration as the proposed algorithm.

2) Evaluation Metrics: In this work, we calculate the following four statistical metrics to evaluate the detection reliability of the proposed TL-LSTM-Ri algorithm, including false-alarm rate (FAR), missing-alarm rate (MAR), fault-detection rate (FDR), and total accuracy (ACC). Concretely, FAR indicates the misclassified negative instances as a percentage of the total negative instances, which is formulated as follows:

$$\mathrm{FAR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}, \tag{15}$$

where FP refers to the instance whose true class is negative while the predicted class is positive; TN represents the instance whose true class is negative and the predicted class is negative. It is important to note that in pipeline fault detection, we uniformly categorize faulty samples as positive and normal samples as negative. In addition, MAR denotes the misclassified positive instances as a percentage of the total positive instances, which is formulated as follows:

$$\mathrm{MAR} = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}}, \tag{16}$$

where FN is the instance whose true class is positive while the predicted class is negative; TP represents the instance whose true class is positive and the predicted class is positive. Moreover, FDR is the correctly classified positive instances as a percentage of the total positive instances, which is formulated as follows:

$$\mathrm{FDR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{17}$$

Furthermore, ACC denotes the correctly classified instances as a percentage of the total instances, which is formulated as follows:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{TN} + \mathrm{FP}}. \tag{18}$$

Besides the above accuracy-related metrics, we also record the detection time (DT) for evaluating the real-time performance of the TL-LSTM-Ri algorithm and all baselines.

3) Experimental Results: Table. I shows the values of FAR, MAR, FDR, ACC, and DT for the TL-LSTM-Ri algorithm and all baselines. It can be found that our TL-LSTM-Ri algorithm achieves the better FAR value (1.00%) than CNN-MMD (-10.3%), DANN (-3.66%), DeepCORAL (-3.16%), and JAN (-1.16%). Moreover, we can also find that the proposed algorithm obtains the best performance in MAR, FDR, and ACC. Therefore, our way of predicting the pipeline data is more efficient than all baselines. In terms of real-time performance, although our algorithm does not achieve the best result, its DT value is small enough to meet the detection requirements. Based on the above experimental results, we can conclude that our TL-LSTM-Ri algorithm can not only satisfy the real-time performance but also further improve the accuracy of fault detection.

V. Conclusion

In this paper, a two-stage incipient fault detection method has been proposed for online PLD where an offline-to-online transfer learning framework has been developed to deal with real-time leakage detection with high false-alarm rate. At the offline training stage, a novel TL-LSTM-Ri has been proposed to transfer the detection knowledge from the PLCs to the PRCs, which has solved the insufficient training data problem. By introducing the regularization terms (i.e., MMD and AB) into the transfer learning framework, the robustness of the PLD model has been enhanced while the false-alarm rate has been reduced. At the online detection stage, a dynamic threshold generated by the proposed TL-LSTM-Ri has been utilized to detect the incipient fault in real time by comparing with the PRCs. Experiment results have shown that the proposed DOTLF outperforms some selected PLD methods under both normal and fault conditions. In the future, we aim to 1) optimize the source domain data to efficiently transfer detection knowledge [10], [19], [48], [55], [57]; and 2) design an advanced control strategy to further improve the detection accuracy and detection rate of the TL-LSTM-Ri-based PLD method [18], [21], [28], [50], [60].

References

[1] S. Azimi, R. Wadhawan, and T. K. Gandhi, Intelligent monitoring of stress induced by water deficiency in plants using deep learning, *IEEE Transactions on Instrumentation and Measurement*, vol. 70, art. no. 5017113, 2021.

[2] S. Bai, J. Z. Kolter, and V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, *arXiv preprint arXiv:1803.01271*, 2018.

[3] U. Baroudi, A. A. Al-Roubaiey, and A. Devendiran, Pipeline leak detection systems and data fusion: A survey, *IEEE Access*, vol. 7, pp. 97426–97439, 2019.

[4] H. Chen, Z. Chai, B. Jiang, and B. Huang, Data-driven fault detection for dynamic systems with performance degradation: A unified transfer learning framework, *IEEE Transactions on Instrumentation and Measurement*, vol. 70, art. no. 3504712, 2020.

[5] K. Cho, B. Van. Merrinboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, *arXiv preprint arXiv:1406.1078*, 2014.

[6] G. K. Dziugaite, D. M. Roy, and Z. Ghahramani, Training generative neural networks via maximum mean discrepancy optimization, *arXiv preprint arXiv:1505.03906*, 2015.

[7] J. Feng, X. Zhang, S. Lu, and F. Yang, A single-stage enhancement-identification framework for pipeline MFL inspection, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, art. no. 3513813, 2022.

[8] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, Domain-adversarial training of neural networks, *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2096–2030, 2016.

[9] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data, *IEEE Transactions on Industrial Electronics*, vol. 66, no. 9, pp. 7316–7325, 2019.

[10] T. Han, C. Liu, R. Wu, and D. Jiang, Deep transfer learning with limited data for machinery fault diagnosis, *Applied Soft Computing*, vol. 103, art. no. 107150, 2021.

[11] T. Han, C. Liu, W. Yang, and D. Jiang, Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application, *ISA Transactions*, vol. 97, pp. 269–281, 2020.

[12] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[13] X. Hu, H. Zhang, D. Ma, and R. Wang, Hierarchical pressure data recovery for pipeline network via generative adversarial networks, *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 1960–1970, 2022.

[14] X. Hu, H. Zhang, D. Ma, R. Wang, and J. Zheng, Minor class-based status detection for pipeline network using enhanced generative adversarial networks, *Neurocomputing*, vol. 424, pp. 71–83, 2021.

TABLE I: Classification Accuracy(%) and Detection Time(s) of Five TL Algorithms for Online PLD Task.

	FAR	MAR	FDR	ACC	DT
CNN-MMD	11.30 ± 2.19	0.50 ± 0.49	99.50 ± 0.49	91.37 ± 1.95	0.0299
DANN	4.66 ± 1.46	0 ± 0.00	100 ± 0.00	96.50 ± 1.27	0.0219
DeepCORAL	4.16 ± 1.38	0 ± 0.00	100 ± 0.00	96.87 ± 1.21	0.0379
JAN	2.16 ± 1.00	0 ± 0.00	100 ± 0.00	98.37 ± 0.88	0.3020
TL-LSTM-Ri(ours)	$\mathbf{1.00 \pm 0.69}$	$\mathbf{0 \pm 0.00}$	$\mathbf{100 \pm 0.00}$	$\mathbf{99.25 \pm 0.60}$	**0.0249**

[15] T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, Wave physics as an analog recurrent neural network, *Science Advances*, vol. 5, no. 12, art. no. 6946, 2019.

[16] J. Jiao, M. Zhao, J. Lin, and K. Liang, Residual joint adaptation adversarial network for intelligent transfer fault diagnosis, *Mechanical Systems and Signal Processing*, vol. 145, art. no. 106962, 2020.

[17] Y. Ju, X. Tian, H. Liu, and L. Ma, Fault detection of networked dynamical systems: A survey of trends and techniques, *International Journal of Systems Science*, vol. 52, no. 16, pp. 3390–3409, 2021.

[18] H. Li, P. Wu, N. Zeng, Y. Liu, and F. E. Alsaadi, A survey on parameter identification, state estimation and data analytics for lateral flow immunoassay: From systems science perspective, *International Journal of Systems Science*, in press, DOI: 10.1080/00207721.2022.2083262.

[19] J. Li, C. Shen, L. Kong, D. Wang, M. Xia, and Z. Zhu, A new adversarial domain generalization network based on class boundary feature detection for bearing fault diagnosis, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, art. no. 2506909, 2022.

[20] X. Li, H. Jiang, M. Niu, and R. Wang, An enhanced selective ensemble deep learning method for rolling bearing fault diagnosis with beetle antennae search algorithm, *Mechanical Systems and Signal Processing*, vol. 142, art. no. 106752, 2020.

[21] X. Li, Q. Song, Z. Zhao, Y. Liu, and F. E. Alsaadi, Optimal control and zero-sum differential game for Hurwicz model considering singular systems with multifactor and uncertainty, *International Journal of Systems Science*, vol. 53, no. 7, pp. 1416–1435, 2022.

[22] X. Li, W. Zhang, N. X. Xu, and Q. Ding, Deep learning-based machinery fault diagnostics with domain adaptation across sensors at different places, *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 6785–6794, 2020.

[23] Y. Li, Y. Ren, H. Zheng, Z. Deng, and S. Wang, A novel cross-domain intelligent fault diagnosis method based on entropy features and transfer learning, *IEEE Transactions on Instrumentation and Measurement*, vol. 70, art. no. 3526314, 2021.

[24] Y. Li, Y. Song, L. Jia, S. Gao, Q. Li, and M. Qiu, Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning, *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2833–2841, 2021.

[25] J. C. W. Lin, Y. Shao, Y. Djenouri, and U. Yun, ASRNN: A recurrent neural network with an attention model for sequence labeling, *Knowledge-Based Systems*, vol. 212, art. no. 106548, 2021.

[26] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, A survey of deep neural network architectures and their applications, *Neurocomputing*, vol. 234, pp. 11–26, 2017.

[27] Y. Liu, Q. Cheng, Y. Gan, Y. Wang, Z. Li, and J. Zhao, Multi-objective optimization of energy consumption in crude oil pipeline transportation system operation based on exergy loss analysis, *Neurocomputing*, vol. 332, pp. 100–110, 2019.

[28] Y. Liu, Z. Wang, L. Zou, J. Hu, and H. Dong, Distributed filtering for complex networks under multiple event-triggered transmissions within node-wise communications, *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 4, pp. 2521–2534, 2022.

[29] M. Long, H. Zhu, J. Wang, and M. I. Jordan, Deep transfer learning with joint adaptation networks, In: *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, 2017, pp. 2208–2217.

[30] W. Lu, Y. Li, Y. Cheng, D. Meng, B. Liang, and P. Zhou, Early fault detection approach with deep architectures, *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1679–1689, 2018.

[31] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, Deep model based domain adaptation for fault diagnosis, *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2296–2305, 2017.

[32] Q. Meng, X. Lang, M. Lin, Z. Cai, H. Zheng, H. Song, and W. Liu, Leak localization of gas pipeline based on the combination of EEMD and cross-spectrum analysis, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, art. no. 9501209, 2021.

[33] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, Domain adaptation via transfer component analysis, *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2011.

[34] S. J. Pan and Q. Yang, A survey on transfer learning, *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[35] T. Pan, J. Chen, J. Xie, Z. Zhou, and S. He, Deep feature generating network: A new method for intelligent fault detection of mechanical systems under class imbalance, *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 6282–6293, 2021.

[36] V. M. Patel, R. Gopalan, R. Li, and R. Chellappa, Visual domain adaptation: A survey of recent advances, *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 53–69, 2015.

[37] L. Santos, A. Haddad, and I. L. dos. Santos, Process safety leading indicators in oil storage and pipelines: Building a panel of indicators, *Chemical Engineering Transactions*, vol. 77, pp. 73–78, 2019.

[38] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review, *IEEE Access*, vol. 7, pp. 122644–122662, 2019.

[39] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, Gearbox fault diagnosis using a deep learning model with limited data sample, *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6263–6271, 2020.

[40] S. Shao, S. McAleer, R. Yan, and P. Baldi, Highly accurate machine fault diagnosis using deep transfer learning, *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2446–2455, 2019.

[41] Y. Shen, Z. Wang, B. Shen, and H. Dong, Outlier-resistant recursive filtering for multi-sensor multi-rate networked systems under weighted try-once-discard protocol, *IEEE Transactions on Cybernetics*, vol. 51, no. 10, pp. 4897–4908, 2021.

[42] W. Sima, H. Zhang, and M. Yang, Edge-loud collaboration detection approach for small-sample imbalanced faults in power lines, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, art. no. 3513711, 2022.

[43] H. Su, Y. Hu, H. R. Karimi, A. Knoll, G. Ferrigno, and E. De. Momi, Improved recurrent neural network-based manipulator control with remote center of motion constraints: Experimental results, *Neural Networks*, vol. 131, pp. 291–299, 2020.

[44] B. Sun and K. Saenko, Deep coral: Correlation alignment for deep domain adaptation, In: *European Conference on Computer Vision*, 2016, pp. 443–450.

[45] Q. Sun, Y. Zhang, B. Lu, and H. Liu, Flow measurement-based self-adaptive line segment clustering model for leakage detection in water distribution networks, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, art. no. 3510213, 2022.

[46] C. Wang, F. Han, Y. Shen, X. Li, and H. Dong, Full-information particle swarm optimizer based on event-triggering strategy and its applications, *Acta Automatica Sinica*, vol. 66, pp. 1–13, 2020.

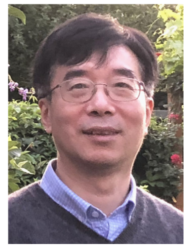
[47] C. Wang, F. Han, Y. Zhang, and J. Lu, An SAE-based resampling SVM ensemble learning paradigm for pipeline leakage detection, *Neurocomputing*, vol. 403, pp. 237–246, 2020.

[48] C. Wang, Z. Wang, Q.-L. Han, F. Han, and H. Dong, Novel leader-follower-based particle swarm optimizer inspired by multiagent systems: Algorithm, experiments, and applications, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, in press, DOI: 10.1109/TSMC.2022.3196853.

[49] F. Wang, W. Lin, Z. Liu, S. Wu, and X. Qiu, Pipeline leak detection by using time-domain statistical features, *IEEE Sensors Journal*, vol. 17, no. 19, pp. 6431–6442, 2017.
[50] L. Wang, S. Liu, Y. Zhang, D. Ding, and X. Yi, Non-fragile l_2-l_∞ state estimation for time-delayed artificial neural networks: An adaptive event-triggered approach, *International Journal of Systems Science*, vol. 53, no. 10, pp. 2247–2259, 2022.
[51] X. Wang and M. S. Ghidaoui, Identification of multiple leaks in pipeline: Linearized model, maximum likelihood, and super-resolution localization, *Mechanical Systems and Signal Processing*, vol. 107, pp. 529–548, 2018.
[52] Y. Wang, Z. Pan, X. Yuan, C. Yang, and W. Gui, A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network, *ISA Transactions*, vol. 96, pp. 457–467, 2020.
[53] L. Wen, L. Gao, and X. Li, A new deep transfer learning based on sparse auto-encoder for fault diagnosis, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 1, pp. 136–144, 2019.
[54] T. Xue, M. Zhong, L. Li, and S. X. Ding, An optimal data-driven approach to distribution independent fault detection, *IEEE Transactions on Industrial Informatics*, vol. 16, no. 11, pp. 6826–6836, 2020.
[55] B. Yang, C. G. Lee, Y. Lei, N. Li, and N. Lu, Deep partial transfer learning network: A method to selectively transfer diagnostic knowledge across related machines, *Mechanical Systems and Signal Processing*, vol. 156, art. no. 107618, 2021.
[56] B. Yang, Y. Lei, F. Jia, and S. Xing, An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings, *Mechanical Systems and Signal Processing*, vol. 122, pp. 692–706, 2019.
[57] B. Yang, Y. Lei, S. Xu, and C. G. Lee, An optimal transport-embedded similarity measure for diagnostic knowledge transferability analytics across machines, *IEEE Transactions on Industrial Electronics*, vol. 69, no. 7, pp. 7372–7382, 2022.
[58] W. Yu and C. Zhao, Low-rank characteristic and temporal correlation analytics for incipient industrial fault detection with missing data, *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 6337–6346, 2021.
[59] H. Zhang, X. Hu, D. Ma, R. Wang, and X. Xie, Insufficient data generative model for pipeline network leak detection using generative adversarial networks, *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 7107–7120, 2022.
[60] J. Zhang, J. Song, J. Li, F. Han, and H. Zhang, Observer-based non-fragile H_∞-consensus control for multi-agent systems under deception attacks, *International Journal of Systems Science*, vol. 52, no. 6, pp. 1223–1236, 2021.
[61] L. Zou, Z. Wang, H. Dong, and Q.-L. Han, Energy-to-peak state estimation with intermittent measurement outliers: The single-output case, *IEEE Transactions on Cybernetics*, vol. 52, no. 11, pp. 11504–11515, 2022.

Chuang Wang received the B.Sc. degree in automation from Northeast Petroleum University, Daqing, China, in 2017, where he is currently pursuing the Ph.D. degree in petroleum and natural gas engineering.

His research interests include evolutionary calculation, deep learning, and transfer learning.



Zidong Wang (Fellow, IEEE) was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics from Suzhou University, Suzhou, China, in 1986, and the M.Sc. degree in applied mathematics and the Ph.D. degree in electrical engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1990 and 1994, respectively.

He is currently a Professor of Dynamical Systems and Computing with the Department of Computer Science, Brunel University London, Uxbridge, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the U.K. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 700 papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, and William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, *Neurocomputing*, and *Systems Science & Control Engineering*, and an Associate Editor for 12 international journals, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, IEEE TRANSACTIONS ON NEURAL NETWORKS, IEEE TRANSACTIONS ON SIGNAL PROCESSING, and IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART C. He is a Member of the Academia Europaea, a Member of the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Weibo Liu received the B.S. degree in electrical engineering from the Department of Electrical Engineering & Electronics, University of Liverpool, Liverpool, U.K, in 2015, and the Ph.D. degree in artificial intelligence from the Department of Computer Science, Brunel University London, Uxbridge, U.K, in 2020.

He is currently a Lecturer in the Department of Computer Science, Brunel University London, Uxbridge, U.K. His research interests include intelligent data analysis, evolutionary computation, transfer learning, machine learning and deep learning. He serves as an Associate Editor for the Journal of Ambient Intelligence and Humanized Computing and the Journal of Cognitive Computation.

Dr. Liu is a very active reviewer for many international journals and conferences.

Yuxuan Shen received the Ph.D. degree in control science and engineering from the Donghua University, Shanghai, China, in 2020.

From June 2018 to September 2018, he was a Research Assistant in the Texas A & M University at Qatar, Doha, Qatar. From November 2018 to November 2019, he was a Visiting Scholar with the Department of Computer Science, Brunel University London, London, U.K. He is currently an Associate Professor with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing, China.

Dr. Shen is a very active reviewer for many international journals.

Hongli Dong (Senior Member, IEEE) received the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2012.

From 2009 to 2010, she was a Research Assistant with the Department of Applied Mathematics, City University of Hong Kong, Hong Kong. From 2010 to 2011, she was a Research Assistant with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From 2011 to 2012, she was a Visiting Scholar with the Department of Information Systems and Computing, Brunel University London, London, U.K. From 2012 to 2014, she was an Alexander von Humboldt Research Fellow with the University of Duisburg-Essen, Duisburg, Germany. She is currently a Professor with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing, China. She is also the Director of the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Daqing. Her current research interests include robust control and networked control systems.

Dr. Dong is a very active reviewer for many international journals.