

Decomposed Degradation Pattern Alignment for Domain Adaptation in Machine Remaining Useful Life Prediction

Byoungmo Koo
Department of Industrial and
Management Engineering
Korea University
Seoul, South Korea
kbn970709@korea.ac.kr

Yunseon Byun
Department of Industrial and
Management Engineering
Korea University
Seoul, South Korea
yun-seon@korea.ac.kr

Jun-Geol Baek*
Department of Industrial and
Management Engineering
Korea University
Seoul, South Korea
jungeol@korea.ac.kr

Abstract—In complex manufacturing processes, multiple sensors have been installed on the machine, enabling the data-driven Remaining Useful Life (RUL) prediction. Timely RUL predictions reduce costs, minimize idle time, and prevent cascading failures in interconnected systems. While high-performance deep learning models have improved RUL prediction accuracy, domain shifts caused by changes in equipment or operating environments often lead to prediction failures. Domain adaptation techniques have been introduced to address this issue, yet they often struggle to align the degradation patterns embedded in run-to-failure data when extracting domain-invariant features. This study presents a novel approach that decomposes time-series sensor data into trend and periodic components, extracts adapted representations for each, and effectively aligns them to improve cross-domain RUL prediction performance. Our proposed method demonstrates superior RUL prediction performance across 12 cross-domain scenarios of the C-MAPSS dataset compared to four baseline models. Notably, our approach proves to be robust and effective in accurately predicting RUL, even in scenarios with significant distributional differences between domains.

Keywords—Domain adaptation, remaining useful life (RUL) prediction, prognostics and health management (PHM), transfer learning, time-series decomposition

I. INTRODUCTION

Prognostics and Health Management (PHM) has emerged as an essential field for optimizing equipment operation and maintenance. The primary objective of PHM is to facilitate predictive maintenance by forecasting the future condition of equipment based on its current operational state [1]. In particular, the prediction of Remaining Useful Life (RUL), which is defined as the time remaining from the current point until failure, is the most extensively studied domain within predictive maintenance. Accurate RUL prediction can effectively minimize lead time, reduce unexpected downtime, and optimize maintenance costs [2]. However, predicting RUL under complex and dynamically changing operating conditions remains a challenging task because RUL is influenced by numerous operational conditions.

Recently, data driven methods, which do not require domain-specific expertise, have demonstrated promising performance in RUL prediction across various industrial environments. Recurrent neural network (RNN) architectures, including Long Short-Term Memory (LSTM) networks [4] and Gated Recurrent Units (GRU) [5], have been effectively applied for RUL prediction capturing temporal dependencies

in sequential sensor data analysis. Transformer-based models have also been employed for their capability to handle complex sequence-to-sequence relationships, with applications such as lithium-ion battery RUL prediction showcasing their strengths in modeling both short-term and long-term dependencies [6]. Despite these efforts, when the target equipment for model application differs or there are discrepancies between the training and testing sensor data environments, models may fail to accurately predict RUL. These challenges are commonly known as domain shift problems.

Domain adaptation methods address domain shift problems by aligning knowledge from existing domains with unseen domains for effective RUL prediction. For instance, domain-adversarial approaches leveraging LSTM architectures have been used to align feature spaces across domains, capturing temporal dependencies [7]. Additionally, contrastive adversarial domain adaptation techniques have been employed to reduce domain discrepancies with InfoNCE loss [8]. However, these approaches cannot consider trend and periodic pattern of degradation and rely on unstable adversarial training, limiting effectiveness in extracting domain-invariant representations for RUL prediction from degraded time-series sensor data.

To address these challenges, we propose the decomposed degradation pattern alignment method. Specifically, we leverage the domain-invariant nature of the trend degradation pattern's inter-sensor associative structure to enhance RUL prediction. Additionally, we capture intervals where periodic patterns are distorted, reflecting the degraded state of the system. The decomposition method for time series is inspired by the Moving Average approach used in Autoformer [9], while the extraction of inter-variable associative structures is motivated by the Sparse Associative Structure Alignment (SASA) framework [10].

The key contributions of the method are as follows:

- Decomposing the equipment's degradation state into trend degradation and periodic pattern distortion enables more accurate RUL prediction.
- Aligning degradation trend features across domains enhances the effectiveness of domain adaptation.
- Achieved superior RUL prediction performance across 12 cross-domain scenarios in the C-MAPSS dataset, demonstrating robust prediction capability even in cases with significant domain distribution differences.

*Corresponding author—Tel: 82-2-3290-3396; Fax: +82-2-3290-4550

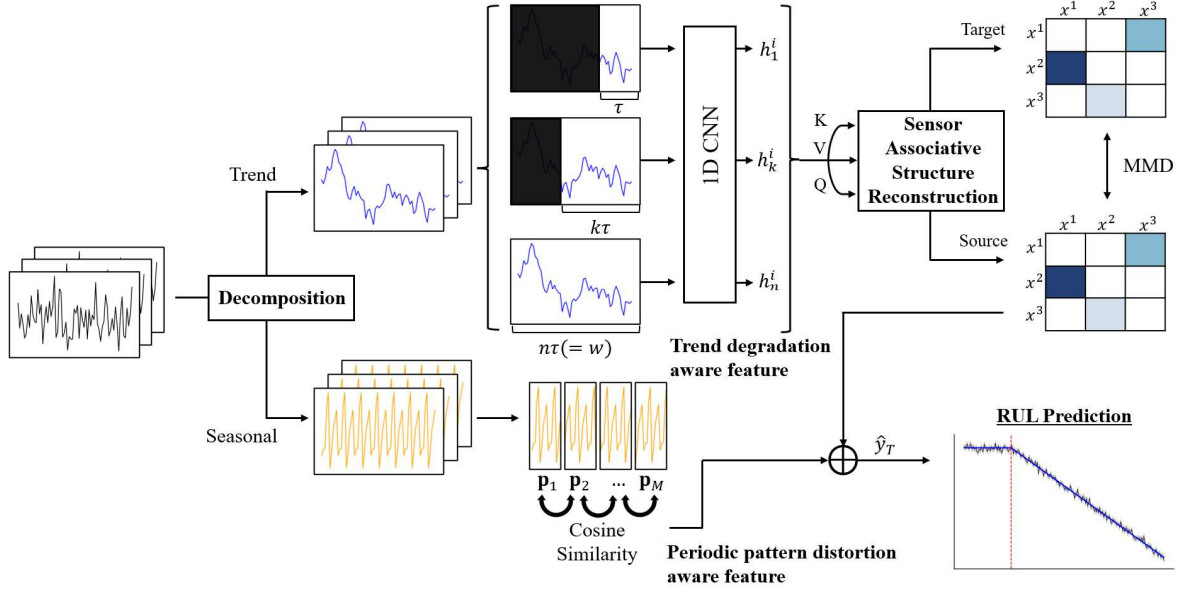


Fig. 1. Overview of our model

II. METHODOLOGY

In this section, we present the proposed methodology, which is divided into two parts: degradation pattern extraction and alignment followed by RUL prediction. The overall framework of the model can be found in Fig. 1.

A. Problem Definition

The sensor data collected for RUL prediction $\mathbf{X} = [x_1, \dots, x_{T-1}, x_T] \in \mathbb{R}^{N \times T}$ is multivariate time series data where N denotes the number of sensor variables, and T represents the total length of the time series. Since the model requires fixed-size time series inputs for training, the sliding window method is employed to transform the data into fixed-size samples $\mathbf{x} = [x_{t-w+1}, \dots, x_{t-1}, x_t]$ where w is the window size. The RUL at the final time point t of the window is set as $y \in \mathbb{R}$. We denote the labeled source domain data as $\mathcal{D}_S = \{\mathbf{x}_S, y_S\}$ and the unlabeled target domain data as $\mathcal{D}_T = \{\mathbf{x}_T\}$. The purpose of this study is to accurately predict y_T by aligning the features extracted during training on $\mathbf{x}_S$ with $\mathbf{x}_T$.

B. Decomposed Degradation Pattern Alignment

1) *Time-series decomposition*: Model captures equipment degradation by decomposing time-series sensor data into trend and seasonal components. The trend component reflects the overall change in sensor values over time, while the seasonal data represents periodic patterns after removing the trend. If degradation originates from the same underlying cause, the trend in sensor values is expected to exhibit similar behavior. Furthermore, abnormal operation caused by degradation may distort the periodic patterns observed during normal operation. By decomposing the data and extracting representations for each component, the degradation state of the equipment can be effectively identified, enabling more accurate predictions of the RUL.

The decomposition process is based on the Series Decomposition Block from the Autoformer [9]. To extract the trend component, padding is added to the beginning of the time-series data, followed by the application of a Moving Average

method. Padding is used to preserve the length of the time series. Subsequently, the seasonal component is obtained by subtracting the extracted trend component from the original time series. It is as follows:

$$\begin{cases} \mathbf{x}_{tr} = \text{AvgPool}(\text{Padding}(\mathbf{x})) \\ \mathbf{x}_{se} = \mathbf{x} - \mathbf{x}_{tr} \end{cases} \quad (1)$$

where $\mathbf{x}_{tr}$ and $\mathbf{x}_{se} \in \mathbb{R}^{N \times T}$ denotes the trend component and seasonal component, respectively.

2) *Trend degradation aware module*: This module is designed for the purpose of capturing the gradual sensor degradation trend over time. Using $\mathbf{x}_{tr}$ as input, candidate segments are generated by gradually increasing the time span from the end of the time series by τ intervals, which can be flexibly adjusted. Additionally, $n\tau$ is set to be equal to w as follows:

$$\begin{aligned} \tilde{\mathbf{x}}_{tr} &= [x_{t-\tau+1:t}, \dots, x_{t-k\tau+1:t}, \dots, x_{t-n\tau+1:t}]_{tr} \\ &= [x_{t-\tau+1:t}, \dots, x_{t-k\tau+1:t}, \dots, x_{t-w+1:t}]_{tr} \end{aligned} \quad (2)$$

Candidate segments are generated for each sensor individually because, even under the same degradation condition, the starting point of degradation may vary for each sensor. Subsequently, a 1D Convolutional neural network (1D-CNN) block with an identical structure is applied to each sensor segment to extract local representations $\mathbf{h}_k$ as follows:

$$\mathbf{h}_k = f(x_{t-k\tau+1:t}; \boldsymbol{\theta}) \quad (3)$$

where $\boldsymbol{\theta}$ denotes the parameter of 1D-CNN block. We can obtain $\mathbf{h} = \{\mathbf{h}_1, \dots, \mathbf{h}_k, \dots, \mathbf{h}_n\}$ by applying to each segment. Finally, the representations of each segment are processed using self-attention to obtain the importance weights of segments for each sensor as follows:

$$\mu_k = \frac{1}{n} \sum_{j=1}^n \frac{(h_k \mathbf{W}^Q)(h_j \mathbf{W}^K)}{\sqrt{d_h}}$$

$$\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_k, \dots, \alpha_n\}$$

$$= \text{softmax}(\mu_1, \mu_2, \dots, \mu_k, \dots, \mu_n) \quad (4)$$

$$Z = \sum_{k=1}^n \alpha_k \cdot (h_k \cdot \mathbf{W}^V)$$

in which $\mathbf{W}^Q, \mathbf{W}^K, \mathbf{W}^V$ are trainable projection parameters and $\sqrt{d_h}$ is the scaling factor.

3) *Periodic pattern distortion aware module*: Periodic patterns embedded in sensor signals can be distorted under abnormal operating conditions. Seasonal component window $\mathbf{x}_{se}$ is divided into multiple patches of size p , represented as follows:

$$\tilde{\mathbf{x}}_{se} = [\mathbf{x}_{1:p}, \mathbf{x}_{p+1:2p}, \dots, \mathbf{x}_{(M-1)p+1:Mp}]_{se}$$

$$= [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M] \quad (5)$$

where $\mathbf{p}_i$ is i -th patch and M is the number of patches. Subsequently, cosine similarity $\mathbf{s}^{ij} = \text{sim}(\mathbf{p}_i, \mathbf{p}_j) = \frac{\mathbf{p}_i \cdot \mathbf{p}_j}{\|\mathbf{p}_i\| \cdot \|\mathbf{p}_j\|}$ is computed for each combination of patches. Under normal operating conditions, cosine similarity remains high and consistent, but under abnormal conditions, it becomes distorted and decreases.

4) *Sensor associative structure alignment*: Using the weights calculated through Trend degradation aware module, the associative structure between sensors is derived. To account for changes in a specific sensor's trend, variations in other sensors exhibit a certain time delay. Furthermore, to emphasize strong inter-sensor relationships and suppress weaker relationships, the *sparsemax* function was employed. Therefore, the associative structure between i th sensor and j th sensor, considering time delay, is defined as follows:

$$e_k^{ij} = \frac{Z^i \cdot h_k^j}{\|Z^i\| \cdot \|h_k^j\|}$$

$$\mathbf{e}^{ij} = \{e_1^{ij}, e_2^{ij}, \dots, e_k^{ij}, \dots, e_n^{ij}\}$$

$$\boldsymbol{\beta}^i = \{\boldsymbol{\beta}^{i1}, \boldsymbol{\beta}^{i2}, \dots, \boldsymbol{\beta}^{ij}, \dots, \boldsymbol{\beta}^{iN}\}$$

$$= \text{sparsemax}(\{\mathbf{e}^{i1}, \mathbf{e}^{i2}, \dots, \mathbf{e}^{ij}, \dots, \mathbf{e}^{iN}\}) \quad (6)$$

The sensor associative structure $\boldsymbol{\beta}$ reflects the trend degradation pattern. Finally, the sensor associative structures from the source and target domains are aligned using the Maximum Mean Discrepancy (MMD) loss function, which can be expressed as follows:

$$\mathcal{L}_{MMD} = \sum_{n=1}^N \left\| \frac{1}{|\mathbf{x}_S|} \sum \boldsymbol{\beta}_S^n - \frac{1}{|\mathbf{x}_T|} \sum \boldsymbol{\beta}_T^n \right\| \quad (7)$$

C. RUL Prediction

The sensor associative structure $\boldsymbol{\beta}$, reflecting the trend degradation pattern, and representing periodic pattern distortion $\mathbf{s}$, are concatenated to form the final feature $\mathbf{U} = \text{concat}(\boldsymbol{\beta}, \mathbf{s})$. The final feature $\mathbf{U}$ is passed through Multi-layer perceptron (MLP) regressor to generate the RUL prediction value, expressed as follows:

$$\hat{y}_{pred} = \text{MLP}(\mathbf{U}; \theta^U) \quad (8)$$

where θ^U denotes the parameters of the MLP regressor. The regression loss function is defined using Mean Squared Error (MSE) as follows:

$$\mathcal{L}_{RUL} = \sum (\hat{y}_{pred} - y_{true}) \quad (9)$$

During training, the MMD loss function, and the regression loss function are jointly optimized. The final training loss $\mathcal{L}$ function is expressed as follows:

$$\mathcal{L} = \mathcal{L}_{RUL} + \gamma \cdot \mathcal{L}_{MMD} \quad (10)$$

where γ denotes hyper-parameter as scaling factor.

III. EXPERIMENTS

A. Dataset

We utilized the C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) benchmark dataset [12] to evaluate the domain adaptation performance of our proposed model. As summarized in Table I, the dataset consists of four subsets: FD001, FD002, FD003, and FD004, which differ in terms of working conditions, fault types, and number of engines. Specifically, FD001 and FD002 were collected under High-Pressure Compressor (HPC) degradation scenarios, while FD003 and FD004 include data from both HPC degradation and Fan degradation scenarios. The sensor data collected for modeling damage propagation consists of measurements from 21 sensors, including temperature, pressure, and speed. Each of the four subsets exhibits distinct variations in sensor data, resulting in differing distributions across the subsets.

TABLE I
DESCRIPTION OF C-MAPSS DATASET

Dataset	FD001	FD002	FD003	FD004
# Training engine units	100	260	100	249
# Testing engine units	100	259	100	248
# Operating conditions	1	6	1	6
# Fault modes	1	1	2	2

B. Baselines

We compared the performance of three domain adaptation for RUL prediction models, ADA-RUL [11], LSTM-DANN [7], and CADA [8], along with the model SASA [10] that motivated our methodology. Since the SASA is designed for general time series domain adaptation tasks, we modified the predictor component into an RUL regressor to adapt it for our experiments. All experiments were conducted five times using different random seeds, and the average performance over these runs was reported.

TABLE II
RESULTS OF EXPERIMENTS

Metric	RMSE					Score				
Method	LSTM-DANN	ADARUL	CADA	SASA	Proposed	LSTM-DANN	ADARUL	CADA	SASA	Proposed
FD001->FD002	56.96	21.10	19.62	<u>18.49</u>	16.90	135,053	4,551	2,174	<u>1,765</u>	1,201
FD001->FD003	41.08	31.07	38.82	<u>18.37</u>	17.95	18,621	5,880	10,571	<u>1,492</u>	947
FD001->FD004	43.26	36.59	31.19	<u>20.01</u>	18.11	20,538	17,037	11,629	<u>2,485</u>	2,379
FD002->FD001	29.52	17.85	<u>13.95</u>	14.42	13.13	15,745	705	<u>344</u>	463	318
FD002->FD003	38.89	32.68	34.09	<u>23.00</u>	19.77	66,547	16,089	5,663	<u>1,442</u>	1,013
FD002->FD004	32.90	32.52	35.39	21.17	<u>22.90</u>	60,778	23,299	19,876	2,835	<u>3,311</u>
FD003->FD001	25.51	34.72	<u>20.67</u>	20.28	21.00	7,095	30,817	<u>1901</u>	1,939	1,664
FD003->FD002	43.99	37.51	<u>21.07</u>	21.28	21.04	94,730	160,628	18,389	4,847	<u>6,023</u>
FD003->FD004	43.61	26.97	21.82	<u>20.81</u>	17.73	137,642	9842	3,868	<u>3,407</u>	2,109
FD004->FD001	27.80	38.14	22.44	<u>19.59</u>	19.25	19,209	71,217	11,164	1,334	<u>1,387</u>
FD004->FD002	26.28	36.41	18.63	<u>17.78</u>	17.60	37,869	50,331	4,583	<u>3,338</u>	3,235
FD004->FD003	31.03	23.08	<u>14.57</u>	14.97	13.75	13,636	1662	967	<u>734</u>	393

C. Evaluation metric

To evaluate the performance of RUL prediction, we use two metrics: Root Mean Squared Error (RMSE) and Score. RMSE measures the difference between the predicted RUL and the actual RUL and is calculated using the following equation:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (11)$$

where $\hat{y}_i$ and y_i denotes the predicted RUL and true RUL, respectively. RMSE metric squares the differences, which means that it equally weights overestimation and underestimation, providing an indication of how close the predictions are to the actual values. In contrast, the Score metric was introduced in the PHM08 Data Challenge [12] and is defined as follows:

$$Score = \begin{cases} \sum_{i=1}^N e^{\left(\frac{y_i - \hat{y}_i}{13} - 1\right)}, & \text{if } (\hat{y}_i < y_i) \\ \sum_{i=1}^N e^{\left(\frac{\hat{y}_i - y_i}{10} - 1\right)}, & \text{if } (\hat{y}_i \geq y_i) \end{cases} \quad (12)$$

The Score metric assigns different penalties to early and late RUL predictions. From a maintenance perspective, early predictions may lead to unnecessary costs, while late predictions can result in critical issues such as machine failure or operator safety risks. Therefore, higher penalties are imposed on late predictions.

D. Result and analysis

Table II summarizes the domain adaptation results of the proposed method compared to comparison models. The evaluation was conducted across 12 scenarios, involving domain adaptation between the four subsets, FD001, FD002, FD003, and FD004, as source and target domains. The bolded numbers represent the highest performance, while the underlined numbers indicate the second-best performance.

The results demonstrate that domain adaptation performance is relatively higher between domains with the same fault type (i.e., FD001, FD002 and FD003, FD004) compared to domains with different fault types. Notably, the proposed method exhibits robust performance even in scenarios with significant distribution differences between source and target domains, outperforming the baseline models. This superior performance can be attributed to the decomposition of sensor data, which allows the model to capture and align domain-invariant trend degradation effectively. Additionally, the method captures domain-specific seasonal degradation, which

further enhances its ability to detect abnormal states, resulting in more accurate RUL predictions.

IV. CONCLUSION

In this paper, we propose a novel domain adaptation approach for RUL prediction by capturing decomposed degradation patterns and aligning sensor associative structures between the source and target domains. This approach effectively captures the domain-invariant relationships between sensors in trend degradation patterns and incorporates the impact of periodic pattern distortion in abnormal state, thereby contributing to improved RUL prediction performance. Our approach is validated on the C-MAPSS dataset across 12 domain adaptation scenarios, demonstrating superior RUL prediction performance. Notably, it exhibits robust results even in cases with significant domain distribution discrepancies.

In future work, we consider three potential improvements. First, we explore methods for aligning periodic pattern distortions across domains under abnormal operating conditions. While the trends in sensor value changes due to performance degradation may be similar across domains, the distorted periodic patterns can vary significantly. This aspect is particularly critical for equipment dominated by vibration data, such as bearings, where extracting domain-invariant representations is essential. Second, we extend domain adaptation experiments in scenarios involving multiple fault types. For example, the N-CMAPSS dataset includes six fault types, and we leverage this dataset for additional experiments to evaluate the robustness and generalizability of our approach. Lastly, we incorporate recent domain adaptation models for RUL prediction as baselines to validate timeliness. Examples include SEA [13], which aligns sensor correlations via graph structures, and MC-DANN [14], which aligns differentiated subspaces across domains.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2022R1A2C2004457). This work was also supported by the BK21 FOUR funded by the Ministry of Education of Korea and Samsung Electronics Co., Ltd. (IO201210-07929-01).

REFERENCES

- [1] P. Nunes, J. Santos, and E. Rocha, "Challenges in predictive maintenance—A review," *CIRP Journal of Manufacturing Science and Technology*, vol. 40, pp. 53–67, 2023.
- [2] B. Tasçı, A. Omar, and S. Ayvaz, "Remaining useful lifetime prediction for predictive maintenance in manufacturing," *Computers & Industrial Engineering*, vol. 184, p. 109566, 2023.

- [3] H. Li *et al.*, “A Review on Physics-Informed Data-Driven Remaining Useful Life Prediction: Challenges and Opportunities,” *Mechanical Systems and Signal Processing*, vol. 209, p. 111120, 2024.
- [4] M. Ma and Z. Mao, “Deep-convolution-based LSTM network for remaining useful life prediction,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1658-1667, Mar. 2020.
- [5] Chen, Zhenghua, et al., “Machine remaining useful life prediction via an attention-based deep learning approach,” *IEEE Transactions on Industrial Electronics*, vol. 68, no. 3, pp. 2521-2531, Mar. 2020.
- [6] Chen, Daoquan, Weicong Hong, and Xiuze Zhou, “Transformer network for remaining useful life prediction of lithium-ion batteries,” *IEEE Access*, vol. 10, pp. 19621-19628, 2022.
- [7] P. R. de O. da Costa, et al., “Remaining useful lifetime prediction via deep domain adaptation,” *Reliability Engineering & System Safety*, vol. 195, p. 106682, 2020.
- [8] M. Ragab *et al.*, “Contrastive Adversarial Domain Adaptation for Machine Remaining Useful Life Prediction,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5239-5249, 2020.
- [9] H. Wu, et al., “Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 22419–22430, 2021.
- [10] R. Cai *et al.*, “Time Series Domain Adaptation via Sparse Associative Structure Alignment,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 8, pp. 6722-6730, 2021.
- [11] M. Ragab, et al., “Adversarial transfer learning for machine remaining useful life prediction,” in *2020 IEEE International Conference on Prognostics and Health Management (ICPHM)*, IEEE, 2020.
- [12] A. Saxena, et al., “Damage propagation modeling for aircraft engine run-to-failure simulation,” in *2008 International Conference on Prognostics and Health Management*, IEEE, 2008.
- [13] Y. Wang, et al., “Sensor alignment for multivariate time-series unsupervised domain adaptation,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 8, 2023.
- [14] S. Xiang, et al., “Micro transfer learning mechanism for cross-domain equipment RUL prediction,” *IEEE Transactions on Automation Science and Engineering*, 2024.