Partial Domain Adaptation for Intelligent Machinery Fault Diagnosis: Leveraging Healthy-Only Target Data for Multi-Class Classification

Hanqi Su[1], Dai-Yan Ji[1], Shinya Tsuruta[2], Daichi Arimizu[2], Yuto Hachiya[2], Koji Wakimoto[2], Jay Lee[1]

[1] *Center for Industrial Artificial Intelligence, Department of Mechanical Engineering,*
University of Maryland, College Park, MD, 20742, USA
hanqisu@umd.edu
jidn@umd.edu
leejay@umd.edu

[2] *Mitsubishi Electric Corporation, Kamakura-shi, Kanagawa, 247-8501, Japan*
Tsuruta.Shinya@dw.MitsubishiElectric.co.jp
Arimizu.Daichi@dp.MitsubishiElectric.co.jp
Hachiya.Yuto@ea.MitsubishiElectric.co.jp
Wakimoto.Koji@df.MitsubishiElectric.co.jp

ABSTRACT

Accurate gearbox fault diagnosis across different operating conditions plays an important role in prognostics and health management. In real industrial scenarios, a common challenge arises when the source domain contains multiple fault classes, while the target domain includes only healthy samples during training. To address this issue, this study proposes a unified industrial fault diagnosis framework designed to handle the partial domain adaptation problem. Specifically, the overall framework involves: a unified data processing pipeline, a robust deep learning architecture for accurate fault classification, and integration of maximum mean discrepancy loss to align feature distributions between source and target domains. Experimental results demonstrate that our proposed partial domain adaptation-based deep learning model significantly outperforms benchmark models, achieving accuracy improvements exceeding 20% across multiple domain adaptation tasks. This study provides a practical solution for intelligent gearbox diagnosis under domain shift constraints.

1. INTRODUCTION

In the prognostics and health management (PHM) domain, the gearbox system often plays an important role in many industrial applications, and effective fault diagnosis is critical for maintaining operational reliability and safety of rotating machinery (Tsui, Chen, Zhou, Hai, & Wang, 2015; Yucesan, Dourado, & Viana, 2021; Polverino et al., 2023; Lee & Su, 2025). Failures in gearbox systems can cause unexpected downtime and high maintenance costs. A major challenge in real-world diagnostics is domain shift, which arises when there are distribution discrepancies between the source domain and the target domain (Farahani, Voghoei, Rasheed, & Arabnia, 2021; Singhal, Walambe, Ramanna, & Kotecha, 2023; S. Zhang et al., 2023). It will significantly affect the performance of diagnostic models when the source and target domains differ substantially. Therefore, many deep learning based domain adaptation methods have been proposed to address this problem in recent years (X. Liu et al., 2022; Yao, Kang, Zhou, Rawa, & Abusorrah, 2023; Yan et al., 2024; Lee, Su, Ji, & Minami, 2025). Despite these advancements, in real scenarios, labeled fault data are often available only in the source domain, while the target domain contains only healthy samples during training, with no labeled fault data. This setting is commonly referred to as a partial domain adaptation problem.

To address these challenges, this study proposes a unified industrial fault diagnosis framework designed for the partial domain adaptation problem. This framework enables accurate multi-class gearbox fault classification across different domains, even when the target domain is limited exclusively to healthy samples during training. The main contributions include an efficient data processing pipeline for feature extraction, a robust convolutional neural network-based feature extractor, and an optimized loss function integrating cross-entropy with maximum mean discrepancy loss to align feature distributions across domains.

The rest of this paper is structured as follows: Section 2 reviews relevant literature, formulates the problem, and provides the dataset description used in this study. Section 3 presents the proposed methodology framework addressing the partial domain adaptation problem. Experimental settings and model results are discussed in Section 4. Finally, Section 5 concludes the paper with key findings and directions for future research.

2. BACKGROUND

2.1. Literature review

Domain adaptation is a subtopic of transfer learning. It aims to address the domain shift problem, where the training (source) and testing (target) data have different distributions (Farahani et al., 2021; Singhal et al., 2023; S. Zhang et al., 2023). This phenomenon is common in real-world industrial applications due to factors such as changes in environmental conditions, sensor drift, and varying working regimes. All of these can cause the underlying data distribution to change. Directly applying a model trained on one domain to a different domain without adaptation often results in significant performance degradation (Shi, Ying, & Yang, 2022; Singhal et al., 2023; Su & Lee, 2024). Meanwhile, collecting labeled data under all possible operating conditions and fault modes is often prohibitively expensive and sometimes impractical. Therefore, there is a need to develop domain adaptation techniques (aiming to align features or distributions across domains) to make the model more robust and generalizable.

Recent advancements in deep learning have significantly enhanced domain adaptation techniques in PHM domain. Methods such as domain-adversarial neural networks (DANN) (Jiang et al., 2024; He, Zhao, Song, Su, & Liu, 2025; Z. Zhu, Chen, & Tang, 2023), Deep CORAL (Qin, Yao, Wang, & Mao, 2021; Kavianpour, Ghorvei, Kavianpour, Ramezani, & Beheshti, 2022; G. Zhang, Zhou, & Cai, 2023), convolutional neural networks (CNNs)-based models (Mao, Zhang, Qiao, & Li, 2022; H. Liu, Chen, Chen, & Gu, 2022; Y. Zhu, Pei, Wang, Xie, & Qian, 2023), long short-term memory (LSTM)-based models (Z. Liu et al., 2023; Li et al., 2024; Lu et al., 2025), and attention-based models (Cui, Wang, Liu, & Pan, 2024; D. Liu, Cui, Wang, & Cheng, 2025; Y. Zhang et al., 2023) have been proposed. For instance, (Y. Zhu et al., 2023) proposed a partial domain adaptation approach for wind turbine gearbox diagnosis. Specifically, they integrated a residual convolutional network enhanced by attention mechanisms to extract discriminative features and developed a weighted adversarial network under a novel weighting strategy. In addition, (Xu, Dai, Zhao, Liu, & He, 2024) introduced a manifold embedded ensemble partial domain adaptation framework, combined with maximum mean and covariance discrepancy and a novel joint weighting mechanism, addressing negative transfer in gearbox fault diagnosis. For cross-machine diagnosis, (Cui et al., 2024) proposed a dictionary domain adaptation transformer (DDAT) for rolling bearing diagnosis. In detail, DDAT utilized a feature dictionary constructed from multi-batch data and was trained under a novel dictionary adaptation framework. A domain-shared transformer architecture using multi-head attention was introduced for learning domain-invariant features. Moreover, (Mao et al., 2022) developed a fusion domain-adaptation CNN using multimodal data (combining vibration signals and infrared thermal images) for gearbox fault diagnosis. Adversarial training was applied to transfer diagnosis knowledge effectively across different operating conditions. In another study, (Kavianpour et al., 2022) proposed a deep coral adversarial network (DCAN), using a CNN backbone for feature extraction and deploying deep coral adaptation and domain adversarial learning to reduce distribution discrepancies between source and target domains. Meanwhile, (Su & Lee, 2025) introduced an ensemble-based transfer learning framework for gearbox fault diagnosis. The framework integrates multiple deep learning architectures, including CNNs, hybrid LSTMs, and transformers, and employs transfer learning to enable adaptation to new data without requiring full retraining. Lastly, (He et al., 2025) developed a deep transfer learning framework for gearbox diagnosis. They utilized convolutional block attention modules and discriminant loss functions to capture distinguishable features. To handle feature distribution inconsistency, they applied the multiple kernel maximum mean discrepancy (MK-MMD) method for global distribution alignment and employed adversarial strategies for sub-domain distribution alignment.

2.2. Problem formulation

This study focuses on a partial domain adaptation (PDA) setting, where the label space of the target domain during training is a subset of that of the source domain. Specifically, the source domain dataset is denoted as $\mathcal{D}_s = \{(x_s^i, y_s^i)\}_{i=1}^{n_s}$, including labeled samples from all three health conditions: healthy, gear wear, and bearing corrosion. The target domain dataset is denoted as $\mathcal{D}_t = \{(x_t^i, y_t^i)\}_{i=1}^{n_t}$. During training, only a subset $\mathcal{D}_t^{\text{train}}$ is used, which includes only a small number of labeled samples from the healthy class and does not contain any samples with gear wear or bearing corrosion class. The remaining portion of the target domain $\mathcal{D}_t^{\text{test}}$ is used strictly for performance evaluation, which contains samples from all three classes. The goal is to learn a classifier f that is trained on $\mathcal{D}_s$ and $\mathcal{D}_t^{\text{train}}$, in order to generalizes well to the entire target domain $\mathcal{D}_t^{\text{test}}$.

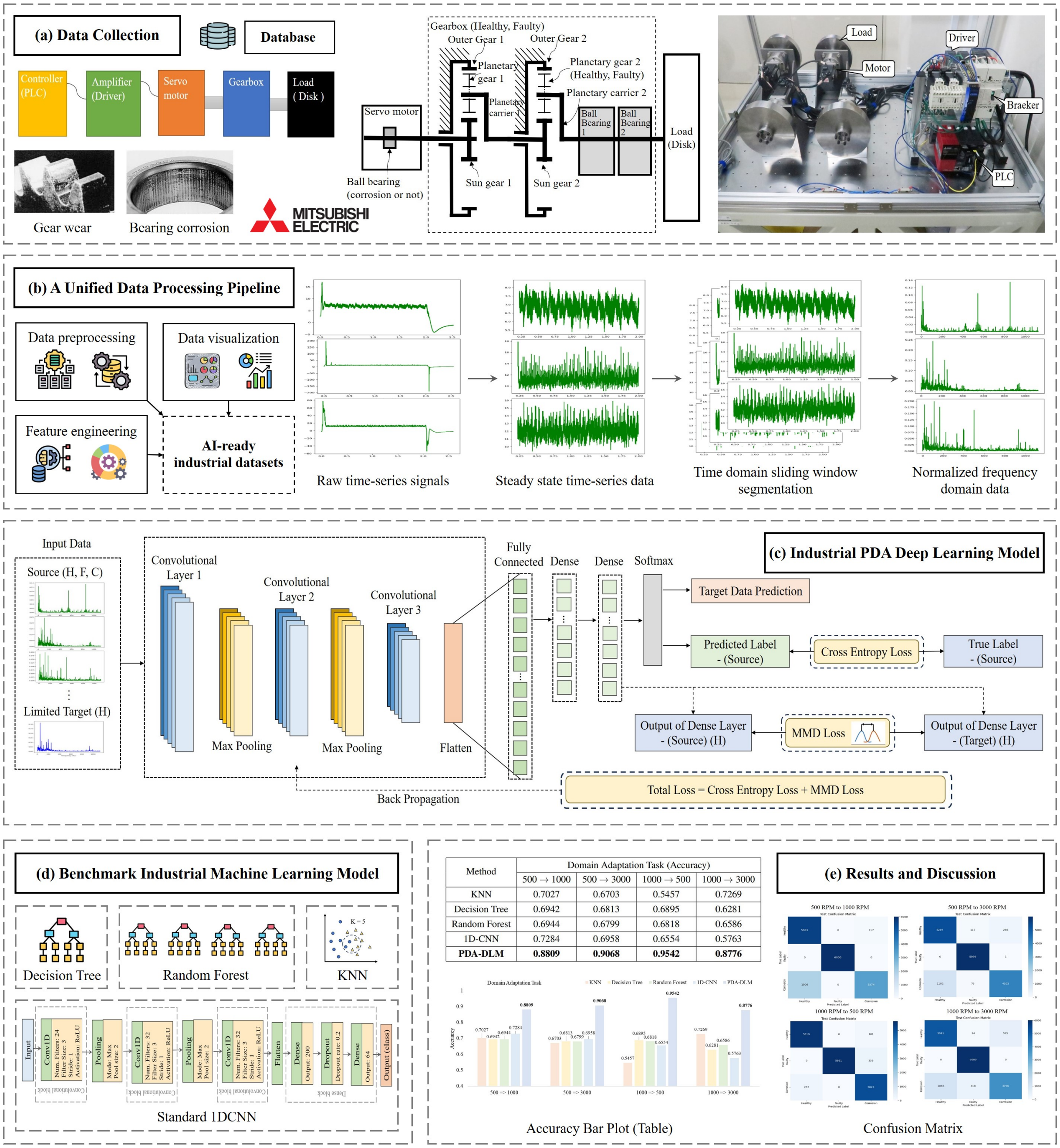

Figure 1. A unified industrial fault diagnosis framework designed for the partial domain adaptation problem, consisting of: (a) data collection; (b) a unified data processing pipeline; (c) an industrial PDA-based deep learning model; (d) benchmark industrial machine learning models; and (e) results and discussion.

2.3. Dataset description

In this study, the dataset is collected from an industrial planetary gearbox system operating under different operational conditions. The experimental setup is illustrated in Figure 1.(a). The dataset comprises three classes: (1) healthy (gear and bearing in normal condition), (2) gear wear (faulty gear and healthy bearing), and (3) bearing corrosion (healthy gear and corroded bearing). Data are collected under three working regimes corresponding to different rotational speeds: 500, 1000, and 3000 RPM. For each regime, 12 distinct command patterns are executed under the high servo gain setting, with a total of 36 patterns across all regimes. Each pattern is repeated 100 times for the healthy and gear wear class and 88 times for the bearing corrosion class. The command patterns include both forward (positive) and backward (negative) rota-

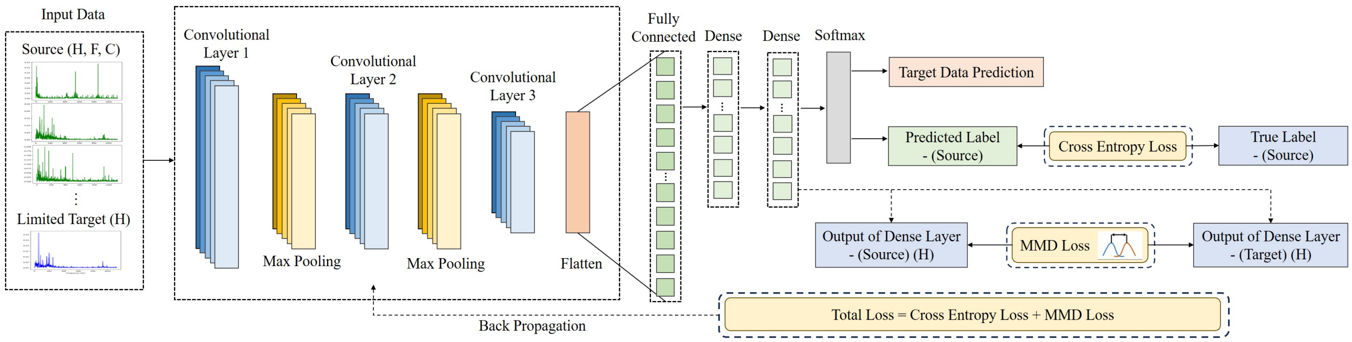

Figure 2. Architecture of the proposed deep learning model for partial domain adaptation.

tions. Moreover, feedback torque current, command velocity, feedback velocity, command position, and feedback position signals are recorded at a fixed sampling interval of 444 μs.

3. METHODOLOGY

Figure 1 presents the overall framework to enhance industrial fault diagnosis under the PDA setting. The process begins with data collection, where a dataset consisting of multiple operating conditions with three classes—healthy (H), gear wear fault (F), and bearing corrosion (C)—is constructed to support this study. A unified data processing pipeline is then applied. It includes data preprocessing, data visualization, and feature engineering to prepare AI-ready industrial datasets. Following this, a PDA-based deep learning model is developed to address the challenge of partial label availability in the target domain. To ensure a comprehensive comparison, multiple benchmark industrial machine learning models are also evaluated.

3.1. Data processing pipeline

To construct a high-quality dataset suitable for model training, a unified data processing pipeline is proposed, as shown in Figure 1.(b). The raw time-series signals contain both transient and steady-state components. Since the steady-state part better reflects stable machine behavior, only this part is retained. After removing the transient components, each signal is truncated to 4000 sampling points. Next, outlier detection is applied to each signal. Abnormal points are identified and replaced with the average of their two nearest neighboring points to smooth abrupt noise. Once clean, steady-state signals are obtained, a sliding window segmentation technique is used to augment the data. Each signal is divided into five overlapping segments using a window size of 2000 points and a step size of 500 points. Finally, each segment is transformed into the frequency domain using Fast Fourier Transform. The resulting frequency-domain features are then normalized within each working regime (i.e., rotational speed). In summary, a total of 17,280 samples were generated and used for training and evaluation of the proposed method.

3.2. PDA-based deep learning model

In this study, a deep learning model, named PDA-DLM, is proposed to address the partial domain adaptation problem. The model architecture is shown in Figure 2. It consists of a convolutional neural network (CNN)-based feature extractor, followed by fully connected (FC) layers, dense layers, and a softmax classifier. In detail, the CNN feature extractor includes three convolutional layers with kernel size 3, stride 2, and filter sizes of 24, 32, and 32, respectively. Two max pooling layers with a pooling size of 2 are applied after the first and second convolutional layers. The extracted low-dimensional features are then flattened and passed through three dense layers with 200, 64, and 3 neurons, respectively. A softmax activation is used to generate class probabilities. To prevent overfitting, dropout regularization with a rate of 0.2 is applied after the first FC layer.

Additionally, the model is trained using a joint loss function that combines the cross-entropy loss L_c and the Maximum Mean Discrepancy (MMD) loss L_{MMD}. The intuition behind the MMD loss is to minimize the statistical distance between the feature distributions of the source and target domains. In practical terms, the model learns to extract features that are not only discriminative for fault classification but also invariant across different operating conditions. Without such alignment, features extracted from the target domain may deviate significantly from those in the source, leading to degraded generalization performance.

L_c is computed using labeled source/target domain samples, which is defined as:

$$L_c = \frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{k=1}^{K} \mathbf{1}[y_s^i = k] \log \hat{p}_s^i(k) \tag{1}$$

where n_s is the number of samples, K is the number of classes, $\hat{p}_s^i(k)$ is the predicted probability of class k for sample i, and $\mathbf{1}[\cdot]$ is the indicator function. L_{MMD} measures the distance between the feature distributions of source and target domain samples (from the healthy class). Let $\{f_s^i\}_{i=1}^{n_s^{(H)}}$

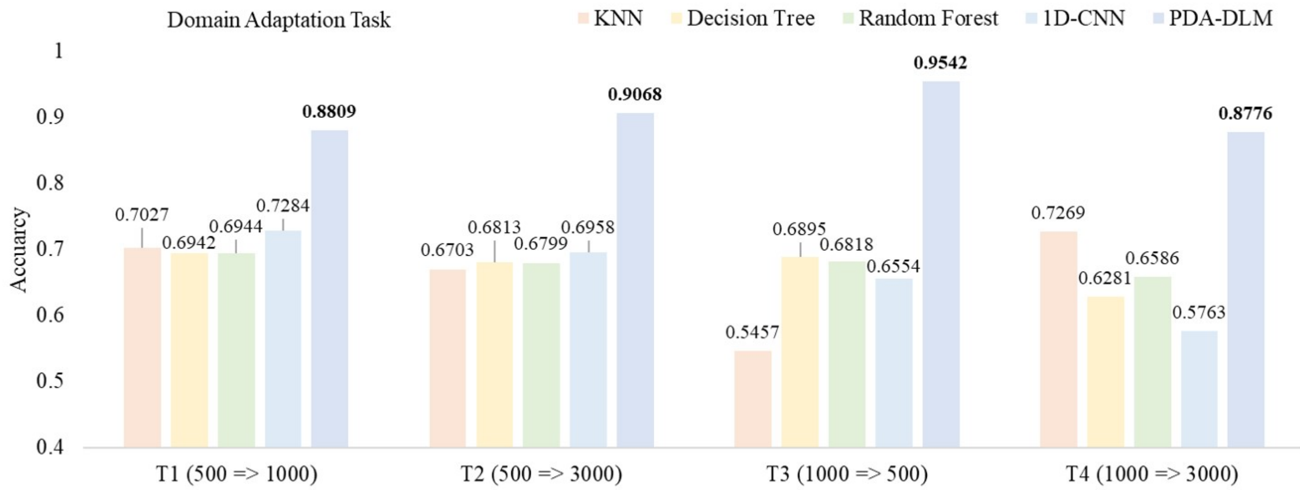

Figure 3. Comparison of classification accuracy across four domain adaptation tasks (T1–T4) using different models: KNN, Decision Tree, Random Forest, 1D-CNN, and the proposed PDA-DLM.

and $\left\{f_t^j\right\}_{j=1}^{n_t^{(H)}}$ denote the features extracted after the second dense layer for healthy samples from the source and target domains, respectively. Then, the MMD loss is computed as:

$$L_{\text{MMD}} = \left\| \frac{1}{n_s^{(H)}} \sum_{i=1}^{n_s^{(H)}} \phi(f_s^i) - \frac{1}{n_t^{(H)}} \sum_{j=1}^{n_t^{(H)}} \phi(f_t^j) \right\|_{\mathcal{H}}^2 \quad (2)$$

where $\phi(\cdot)$ maps features to a reproducing kernel Hilbert space (RKHS), and $\|\cdot\|_{\mathcal{H}}$ denotes the RKHS norm. Here, RKHS is important because it could measure differences between probability distributions in a richer feature space than the raw input. By using kernel functions to map features into RKHS, the model can capture not only mean shifts but also higher-order discrepancies between source and target distributions. This means that even when the two domains appear similar in the original feature space, small differences may still be revealed in RKHS and minimized through MMD, leading to more robust domain alignment.

Finally, the total loss function is defined as:

$$L_{\text{tot}} = \alpha L_c + \beta L_{\text{MMD}} \quad (3)$$

where $\alpha > 0$ and $\beta > 0$ are weighting coefficients that balance the contribution of the two loss components. This ensures the model to learn domain-invariant features while maintaining classification performance on labeled data.

3.3. Benchmark machine learning models

To evaluate the effectiveness of the proposed PDA-DLM model, several benchmark machine learning models are implemented for comparison. These include classical models such as decision tree, random forest, K-nearest neighbors (KNN), and a standard 1D-CNN model sharing a similar architecture to PDA-DLM but excludes MMD loss. All models are trained only using the labeled source domain data. On the one hand, the normalized frequency-domain representations are used as input for 1D-CNN model. On the other hand, for other classical models, principal component analysis (PCA) is applied to the input features, where top 100 principal components are kept.

4. RESULT AND DISCUSSION

4.1. Experimental tasks and hyperparameter settings

In order to evaluate model performance under the partial domain adaptation setting, four domain adaptation tasks are designed using different source and target domains defined by rotational speed (RPM) differences. The tasks are summarized in Table 1. Specifically, tasks T1 and T3 reverse the source and target domains. Differences in performance between these tasks reflect the asymmetric nature of domain adaptation, where transferring knowledge from one domain to another may not be equally effective in reverse. Additionally, tasks T1 and T2 share the same source domain (500 rpm) but target different domains (1000 rpm and 3000 rpm), which helps to analyze adaptation performance under increasing domain discrepancy.

Moreover, the hyperparameter settings for the proposed PDA-DLM model are shown below. The model is trained using a batch size of 128. Early stopping is used with a patience of 8 epochs, and if the validation loss does not improve, the training process will be terminated. In addition, the initial

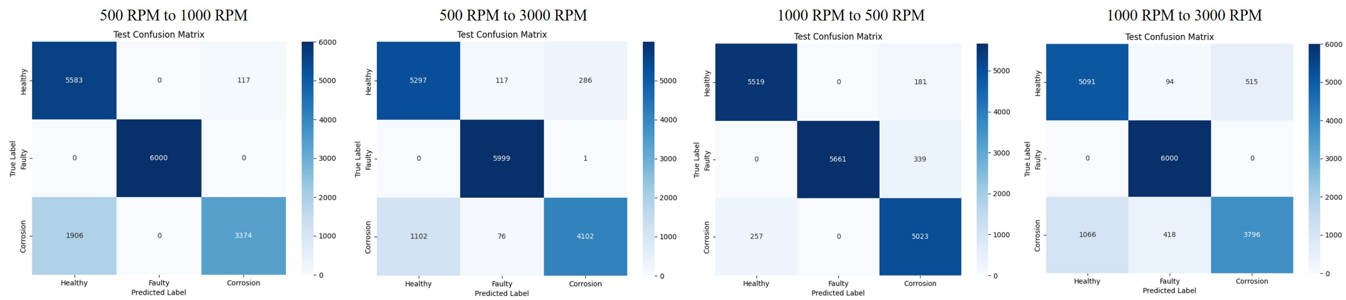

Figure 4. Confusion matrices of PDA-DLM evaluated on the test data on different domain adaptation tasks (best accuracy).

learning rate is set to 0.0008 and optimized using the Adam optimizer. A learning rate scheduler is applied using a step decay strategy, where the learning rate is multiplied by a factor of 0.96 every 5 epochs. Finally, the weighting coefficients for the loss function in PDA-DLM are set as $\alpha = 1.0$ for the cross-entropy loss and $\beta = 0.1$ for the MMD loss. In this study, source and target domains each contain a total of 17,280 samples, including 6000 healthy samples, 6000 gearbox fault (wear) samples, and 5280 bearing corrosion samples. In all experiments, 5% of the healthy samples from the target domain are merged into the source domain during training to simulate the partial domain adaptation setting. The remaining target domain samples are used for final testing. For the training set, 90% of the data is used for training and 10% for validation. This division is under random sampling with a fixed random seed to ensure reproducibility.

Table 1. Description of the Designed Experiments.

Domain Adaptation Task	**Source Domain**	**Target Domain**
T1	500 rpm	1000 rpm
T2	500 rpm	3000 rpm
T3	1000 rpm	500 rpm
T4	1000 rpm	3000 rpm

4.2. Model performance

The classification accuracy results for all compared models are summarized in Table 2 and visualized in Figure 3. The proposed PDA-DLM consistently outperforms all baselines (KNN, Decision Tree, Random Forest, and 1D-CNN) across all domain adaptation tasks, achieving improvements of more than 20% in accuracy in every task. Specifically, in T3 (1000 $\rightarrow$ 500 rpm), PDA-DLM achieves the highest overall accuracy of 95.42%, showing that the model can transfer knowledge effectively even when adapting from high-speed to low-speed conditions. Moreover, in T2 (500 $\rightarrow$ 3000 rpm), where the target domain shifts to a much higher working regime, PDA-DLM still maintains a high accuracy of 90.68%. These results demonstrate that the proposed approach can effectively learn robust and transferable representations under partial domain adaptation settings. In contrast, the classical and 1D-CNN models trained solely on the source domain without domain adaptation show lower and more variable accuracy. For instance, in T3, the accuracy of KNN drops to 54.57%, while PDA-DLM still remains high accuracy. This performance gap highlights the benefit of incorporating MMD loss to address domain adaptation problems.

Table 2. Domain adaptation task result.

Method	Domain Adaptation Task (Accuracy)			
	T1	T2	T3	T4
KNN	0.7027	0.6703	0.5457	0.7269
Decision Tree	0.6942	0.6813	0.6895	0.6281
Random Forest	0.6944	0.6799	0.6818	0.6586
1D-CNN	0.7284	0.6958	0.6554	0.5763
PDA-DLM	**0.8809**	**0.9068**	**0.9542**	**0.8776**

To further understand class-level prediction behavior, confusion matrices for PDA-DLM are shown in Figure 4. Across all tasks, the model demonstrates strong classification performance for the gearbox fault (wear) class, with nearly perfect prediction. Misclassification mainly occurs between healthy samples and corrosion samples. The possible reason lies in their similar frequency-domain patterns within certain speed ranges. For instance, in T1, 1,906 corrosion samples are wrongly classified as healthy samples. In T2 and T4, due to the greater rotational speed difference between the source domain and the target, it might lead to more separable representations, thus reducing confusion.

Beyond technical performance, the proposed PDA-DLM framework has direct implications for industrial PHM. First, robust transferability across different speeds and working regimes reduces the need for extensive fault data collection in every new environment, lowering both cost and effort in model deployment. Second, higher diagnostic accuracy translates into fewer false alarms and missed detections, which can significantly reduce unscheduled downtime. Third, the ability to adapt with only healthy target data aligns well with realistic industrial scenarios, where faulty data are rarely available in advance. Together, these benefits highlight the potential of PDA-DLM to support more reliable, cost-effective, and scalable PHM solutions in practice.

5. CONCLUSION

In this study, we address the PDA problem for gearbox fault diagnosis under different operating conditions. Specifically, we focus on scenarios in which the target domain contains only a limited number of healthy samples during training. The proposed approach integrates deep learning models and domain adaptation learning using MMD, effectively aligning the feature distributions of source and target domains. From a practical perspective, this framework reduces the reliance on labeled fault data in new domains and enables more efficient transfer of diagnostic models to diverse industrial conditions. Experimental results demonstrate that the proposed PDA-DLM significantly outperformed traditional machine learning models and standard CNN baselines, achieving over 20% improvement in classification accuracy across multiple domain adaptation tasks. Despite its strong overall performance, misclassifications between healthy and corrosion classes in some conditions indicate potential areas for further improvement. Future work may explore attention-based mechanisms or transformer-based approaches, adaptive weighting strategies, or the integration of large language models (Lee & Su, 2024).

ACKNOWLEDGMENT

We would like to express our gratitude to our colleague, Naoto Takano, Mitsubishi Electric Corporation, who provided us with gearbox data and many valuable insights about motion systems.

REFERENCES

Cui, L., Wang, G., Liu, D., & Pan, X. (2024). Dictionary domain adaptation transformer for cross-machine fault diagnosis of rolling bearings. *Engineering Applications of Artificial Intelligence*, *138*, 109261.

Farahani, A., Voghoei, S., Rasheed, K., & Arabnia, H. R. (2021). A brief review of domain adaptation. *Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE 2020*, 877–894.

He, X., Zhao, F., Song, N., Su, C., & Liu, P. (2025). A deep-transfer-learning fault diagnosis method for gearboxes based on discriminative feature extraction and improved domain adversarial neural networks. *Nondestructive Testing and Evaluation*, 1–22.

Jiang, F., Lin, W., Wu, Z., Zhang, S., Chen, Z., & Li, W. (2024). Fault diagnosis of gearbox driven by vibration response mechanism and enhanced unsupervised domain adaptation. *Advanced Engineering Informatics*, *61*, 102460.

Kavianpour, M., Ghorvei, M., Kavianpour, P., Ramezani, A., & Beheshti, M. T. (2022). An intelligent gearbox fault diagnosis under different operating conditions using adversarial domain adaptation. In *2022 8th international conference on control, instrumentation and automation (iccia)* (pp. 1–6).

Lee, J., & Su, H. (2024). A unified industrial large knowledge model framework in industry 4.0 and smart manufacturing. *International Journal of AI for Materials and Design*, *1*(2), 41–47.

Lee, J., & Su, H. (2025). Rethinking industrial artificial intelligence: A unified foundation framework. *International Journal of AI for Materials and Design*, *2*(2), 56–68.

Lee, J., Su, H., Ji, D.-Y., & Minami, T. (2025). Engineering artificial intelligence: framework, challenges, and future direction. *Machine Learning: Engineering*, *1*(1), 013001. doi: 10.1088/3049-4761/adce0d

Li, F., Dai, Z., Jiang, L., Song, C., Zhong, C., & Chen, Y. (2024). Prediction of the remaining useful life of bearings through cnn-bi-lstm-based domain adaptation model. *Sensors*, *24*(21), 6906.

Liu, D., Cui, L., Wang, G., & Cheng, W. (2025). Interpretable domain adaptation transformer: a transfer learning method for fault diagnosis of rotating machinery. *Structural Health Monitoring*, *24*(2), 1187–1200.

Liu, H., Chen, W., Chen, W., & Gu, Y. (2022). A cnn-lstm-based domain adaptation model for remaining useful life prediction. *Measurement Science and Technology*, *33*(11), 115118.

Liu, X., Yoo, C., Xing, F., Oh, H., El Fakhri, G., Kang, J.-W., ... others (2022). Deep unsupervised domain adaptation: A review of recent advances and perspectives. *APSIPA Transactions on Signal and Information Processing*, *11*(1).

Liu, Z., Tan, C., Liu, Y., Li, H., Cui, B., & Zhang, X. (2023). A study of a domain-adaptive lstm-dnn-based method for remaining useful life prediction of planetary gearbox. *Processes*, *11*(7), 2002.

Lu, X., Yao, X., Jiang, Q., Shen, Y., Xu, F., & Zhu, Q. (2025). Remaining useful life prediction model of cross-domain rolling bearing via dynamic hybrid domain adaptation and attention contrastive learning. *Computers in Industry*, *164*, 104172.

Mao, G., Zhang, Z., Qiao, B., & Li, Y. (2022). Fusion domain-adaptation cnn driven by images and vibration signals for fault diagnosis of gearbox cross-working conditions. *Entropy*, *24*(1), 119.

Polverino, L., Abbate, R., Manco, P., Perfetto, D., Caputo, F., Macchiaroli, R., & Caterino, M. (2023). Machine learning for prognostics and health management of industrial mechanical systems and equipment: A systematic literature review. *International*

Journal of Engineering Business Management, *15*, 18479790231186848.

Qin, Y., Yao, Q., Wang, Y., & Mao, Y. (2021). Parameter sharing adversarial domain adaptation networks for fault transfer diagnosis of planetary gearboxes. *Mechanical Systems and Signal Processing*, *160*, 107936.

Shi, Y., Ying, X., & Yang, J. (2022). Deep unsupervised domain adaptation with time series sensor data: A survey. *Sensors*, *22*(15), 5507.

Singhal, P., Walambe, R., Ramanna, S., & Kotecha, K. (2023). Domain adaptation: challenges, methods, datasets, and applications. *IEEE access*, *11*, 6973–7020.

Su, H., & Lee, J. (2024). Machine learning approaches for diagnostics and prognostics of industrial systems using open source data from phm data challenges: A review. *International Journal of Prognostics and Health Management*, *15*(2).

Su, H., & Lee, J. (2025). Transfer learning and ensemble learning for fault diagnosis using vibration signals. In *2025 ieee international conference on prognostics and health management (icphm)* (pp. 1–9).

Tsui, K. L., Chen, N., Zhou, Q., Hai, Y., & Wang, W. (2015). Prognostics and health management: A review on data driven approaches. *Mathematical Problems in Engineering*, *2015*(1), 793161.

Xu, H., Dai, L., Zhao, P., Liu, X., & He, C. (2024). Manifold embedded ensemble partial domain adaptation: A novel partial-set transfer mechanism for gearboxes fault diagnosis. *IEEE Sensors Journal*.

Yan, R., Li, W., Lu, S., Xia, M., Chen, Z., Zhou, Z., … Lu, J. (2024). Transfer learning for prognostics and health management: Advances, challenges, and opportunities. *Journal of Dynamics, Monitoring and Diagnostics*, 60–82.

Yao, S., Kang, Q., Zhou, M., Rawa, M. J., & Abusorrah, A. (2023). A survey of transfer learning for machinery diagnostics and prognostics. *Artificial Intelligence Review*, *56*(4), 2871–2922.

Yucesan, Y. A., Dourado, A., & Viana, F. A. (2021). A survey of modeling for prognosis and health management of industrial equipment. *Advanced Engineering Informatics*, *50*, 101404.

Zhang, G., Zhou, T., & Cai, Y. (2023). Coral-based domain adaptation algorithm for improving the applicability of machine learning models in detecting motor bearing failures. *Journal of Computational Methods in Engineering Applications*, 1–17.

Zhang, S., Lei, S., Jiefei, G., Ke, L., Lang, Z., & Pecht, M. (2023). Rotating machinery fault detection and diagnosis based on deep domain adaptation: A survey. *Chinese Journal of Aeronautics*, *36*(1), 45–74.

Zhang, Y., Ji, J., Ren, Z., Ni, Q., Gu, F., Feng, K., … Liu, Z. (2023). Digital twin-driven partial domain adaptation network for intelligent fault diagnosis of rolling bearing. *Reliability Engineering & System Safety*, *234*, 109186.

Zhu, Y., Pei, Y., Wang, A., Xie, B., & Qian, Z. (2023). A partial domain adaptation scheme based on weighted adversarial nets with improved cbam for fault diagnosis of wind turbine gearbox. *Engineering Applications of Artificial Intelligence*, *125*, 106674.

Zhu, Z., Chen, G., & Tang, G. (2023). Domain adaptation with multi-adversarial learning for open-set cross-domain intelligent bearing fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–11.

BIOGRAPHIES

Hanqi Su received his B.Eng. degree in Robotics Engineering from Southern University of Science and Technology, Shenzhen, Guangdong, China, in 2023. Previously, he participated in a one-year Special Student Program at the Massachusetts Institute of Technology (MIT) during his fourth undergraduate year. He is currently pursuing his Ph.D. in Mechanical Engineering at the University of Maryland, College Park. His research is primarily focused on machine learning, deep learning, prognostics and health management, engineering design, and industrial artificial intelligence.

Dai-Yan Ji received his B.S. degree in Electronic Engineering and M.S. degree in Communications Engineering from Feng Chia University, Taiwan, in 2009 and 2012, respectively. He is currently pursuing a Ph.D. degree in Mechanical Engineering at the University of Maryland, College Park, MD, USA. His research interests include machine learning, signal processing, industrial artificial intelligence, and prognostics and health management.

Shinya Tsuruta received his M.S. degree in precision engineering from the University of Tokyo in 2011. From 2011 to 2019, He started his career as a camera hardware engineer and then contributed as an embedded software engineer to improve the performance of industrial printers and develop fault prediction system. Since 2020, He has joined the R&D center at Mitsubishi Electric Corp., Japan, focusing on data analysis and machine learning. His current interests include anomaly detection, conditioned-based maintenance, and the other application of machine learning in manufacturing.

Daichi Arimizu received his B.S. and M.S. degrees from Kyoto University in 2020 and 2022 respectively. His master's research was in hadron physics and he conducted the development of semiconductor-based electron and positron detectors. Since 2022 he has been with Mitsubishi Electric Corp., Japan. His current research interests include data analysis and physics-informed AI.

Yuto Hachiya received his B.S. degree from Tokyo University of Science in 2019 and his M.S. degree from Tohoku University in 2021. Since 2021, he has been with Mitsubishi Electric Corp., Japan, focusing on the digitalization of production sites from 2021 to 2023. Since 2024, he has been engaged in the development of technologies for detecting anomalies in machinery. His research interests include data analysis, anomaly detection, and the estimation of remaining useful life.

Koji Wakimoto received his B.S. and M.S. degrees from Kyoto University in 1983 and 1985 respectively and received his Dr. Eng. degree from the University of Tokyo in 1996. Since 1985 he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis, computer vision and geospatial information systems. He is a member of the IEEE.

Jay Lee received his B.S. degree in mechanical engineering from Tamkang University, Taipei City, Taiwan, in 1979, M.S. degree in mechanical engineering from the University of Wisconsin-Madison, Madison, WI, USA in 1983, M.S. degree in industrial management from the State University of New York-Stony Brook, Stony Brook, NY, USA in 1987, and D.Sc. degree in mechanical engineering from George Washington University, Washington, DC, USA in 1992. Currently, he is a Clark Distinguished Chair Professor in Mechanical Engineering at the University of Maryland, College Park, and the founding director of the Industrial Artificial Intelligence Center to tackle the opportunities and unmet needs in the application of Artificial Intelligence to a wide range of industrial applications. Previously, he served as the Founding Director of the NSF I/UCRC on Intelligent Maintenance Systems (IMS) which has developed research memberships with over 100 global companies since 2000 and was selected as the most economically impactful I/UCRC in the NSF Economic Impact Study Report in 2012. He advised his students to win 1st place 5 times in the PHM Data Challenge during 2008-2016 as well as 1st Place in the PHM AP Data Challenge in 2023.