

Toward Foundation Models for Domain Transfer under Limited Target Data in Prognostics and Health Management

Hanqi Su¹, Jay Lee²

^{1,2} *Center for Industrial Artificial Intelligence, Department of Mechanical Engineering,
University of Maryland, College Park, MD, 20742, USA*
hanqisu@umd.edu, leejay@umd.edu

ABSTRACT

Domain shift remains a major challenge in prognostics and health management (PHM), where machine learning models trained on one operating condition or machine often experience significant performance degradation when transferred to a different target domain. Although domain adaptation methods have been widely studied for intelligent fault diagnosis and related PHM tasks, many existing approaches still rely on sufficient target-domain data. However, target-domain data are often extremely limited when new machines, sensors, or working conditions are encountered in real industrial applications. This research aims to achieve reliable domain transfer with limited target-domain data and to explore the development of pre-trained and foundation-model frameworks for PHM applications. The proposed research follows a progressive path from conventional deep domain adaptation methods to pre-trained transformer-based models and ultimately toward PHM-oriented foundation models capable of learning transferable representations across heterogeneous industrial datasets. The proposed framework integrates time-series tokenization, transformer-based representation learning, large-scale pre-training, and parameter-efficient downstream adaptation for PHM tasks. Overall, foundation-model-based approaches can provide a promising direction for developing more transferable and data-efficient PHM systems.

1. MOTIVATION AND PROBLEM STATEMENT

In the prognostics and health management (PHM) domain, machine learning (ML) models are often developed under the assumption that the training and testing data follow similar distributions. However, in practical industrial settings, this assumption is rarely satisfied. Variations in machines, operating conditions, sensor configurations, and environments often cause a significant domain shift between the source and target data. Consequently, a model trained on abundant labeled data

from one source domain may suffer significant performance degradation when directly transferred to a different target domain (Wang et al., 2025; Su & Lee, 2024; Yan et al., 2024).

A common solution to address domain shift in PHM is to use domain adaptation-based approaches to transfer knowledge from a labeled source domain to a target domain. Existing studies have investigated a variety of settings, including supervised, semi-supervised, and unsupervised domain adaptation (Wang et al., 2025), and many ML approaches have been developed by integrating deep neural networks with techniques such as feature alignment, latent-space alignment, adversarial learning, metric learning, pseudo-labeling, contrastive learning, and reconstruction-based learning (Farahani, Voghoei, Rasheed, & Arabnia, 2021; Shi, Ying, & Yang, 2022; Zhang et al., 2023; Yao, Kang, Zhou, Rawa, & Abusorrah, 2023). These approaches have demonstrated that domain knowledge transfer can significantly improve fault diagnosis performance when sufficient target-domain data are available to support adaptation.

Despite this progress, several important gaps remain. First, much of the existing PHM domain adaptation literature still assumes access to sufficient target-domain data. In practical PHM applications, however, only a very limited amount of target-domain data may be available when a new machine, sensor, or operating condition is encountered (Wang et al., 2025). Second, many existing PHM methods are developed and evaluated on individual datasets or specific operating conditions, which may limit their robustness and transferability across different domains. As the domain shift becomes more severe, these task-specific models may experience substantial performance degradation. Third, the development of transferable large-scale representations for PHM remains limited compared with other fields such as natural language processing and computer vision. While pre-trained and foundation models have demonstrated strong transferability by learning general patterns from large-scale data and adapting them to downstream tasks with limited supervision (Zhou et al., 2025), their application to domain transfer in PHM remains

Hanqi Su et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

in its early stages, particularly for time-series data analytics. Therefore, the core challenge addressed in this thesis is the following: **how can reliable domain transfer be achieved in PHM when the target domain contains only limited data?**

To address this challenge, this thesis proposes a progressive research path for reliable domain transfer under limited target-domain data in PHM. The first stage studies conventional deep domain adaptation methods for PHM. The second stage investigates how pre-trained models can improve cross-domain transfer under limited target data. The final stage aims to move toward PHM-oriented foundation models capable of learning general representations across multiple machines, datasets, and operating conditions, with the long-term goal of enabling more reliable and data-efficient PHM systems.

2. EXPECTED NOVEL CONTRIBUTIONS

This thesis is expected to make the following contributions to the PHM field:

1. Establish a practical domain transfer framework for PHM under severe target-data scarcity. Investigate how limited target-domain data influences cross-domain transfer performance across different machines, operating conditions, and datasets, and provide design principles for robust PHM domain adaptation.
2. Develop a pre-trained model-based domain adaptation methodology that adapts existing large pre-trained models through specialized tokenization and lightweight adaptation strategies for time-series data analysis. It aims to improve transferable representation learning and cross-domain robustness compared with conventional task-specific PHM models.
3. Develop a PHM-oriented foundation model framework capable of learning general and transferable representations across multiple datasets through large-scale pre-training and efficient fine-tuning for downstream adaptation. This framework aims to support more reliable and reusable cross-domain transfer with minimal target-domain data and provide a path toward scalable PHM intelligence.

3. PROPOSED RESEARCH PLAN

3.1. Proposed Methodology

This research focuses on developing pre-trained and foundation-model frameworks for reliable domain transfer with limited target-domain data. As illustrated in Figure 1, the proposed framework integrates heterogeneous industrial datasets, transformer-based time-series tokenization, existing pre-trained backbone models, PHM-oriented large-scale

foundation model pre-training, and downstream adaptation for industrial PHM tasks.

In the proposed framework, raw industrial sensor signals will be collected from multiple PHM datasets that may originate from the same industrial domain (e.g., different bearing or gearbox datasets) or from different domains (e.g., bearing systems, gearbox systems, turbofan engines, and specific manufacturing processes). These datasets may contain different fault categories, operating conditions, sensor configurations, noise levels, and dataset scales, resulting in significant domain shift and data heterogeneity across datasets. In addition, while some source-domain datasets may contain abundant labeled samples, the target-domain dataset may contain only a limited number of samples.

1) Time-series tokenization process: To make industrial sensor data suitable for transformer-based pre-trained model and foundation model training, the raw time-series signals will first be transformed into tokenized representations. Similar to tokenization strategies used in natural language processing and computer vision, the original sensor signals will be segmented into temporal patches through patch-based segmentation. The segmented patches will then be converted into embedding vectors. This tokenization process enables industrial time-series data to be represented as sequential tokens for large-scale representation learning and cross-domain knowledge transfer.

Two technical paths will be investigated in this research.

2-a) Existing pre-trained models for domain transfer: The first technical path investigates the use of existing pre-trained models for PHM domain transfer. Representative backbone models may include general transformer-based models, such as BERT, GPT, T5, and vision transformer (ViT), as well as recent time-series pre-trained models, such as MOMENT and UniTS. These models provide pre-trained weights and have the potential to be adapted to industrial sensor data through appropriate time-series tokenization and downstream fine-tuning strategies. This path aims to examine whether existing pre-trained backbone models can improve domain transfer performance compared with task-specific PHM models trained from scratch.

2-b) Toward PHM-oriented foundation models: The second technical path aims to develop PHM-oriented foundation models through large-scale pre-training on heterogeneous industrial datasets. Unlike existing pre-trained models primarily developed on natural language, vision, or general time-series data, the PHM-oriented foundation model will be trained on industrial PHM datasets. This design aims to learn industrial-specific representations that are more suitable for the PHM domain transfer. To improve cross-domain generalization, different pre-training strategies will be investigated, including masked reconstruction, contrastive learning,

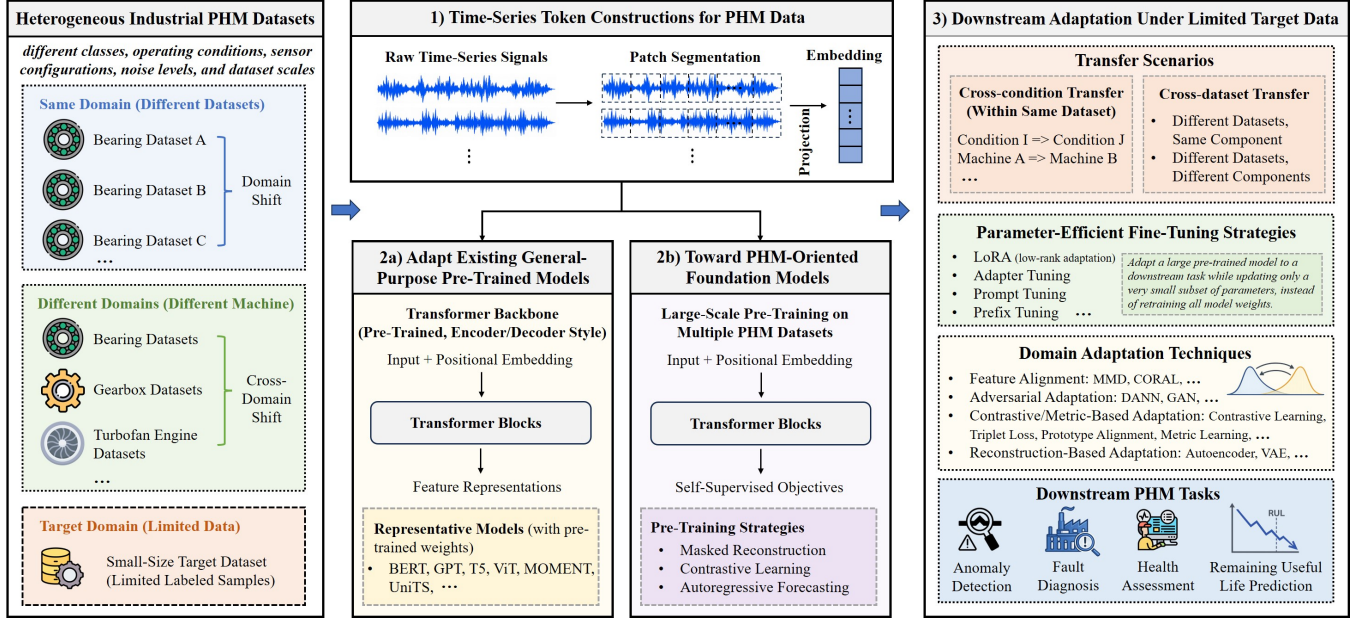


Figure 1. PHM-oriented foundation model framework for domain transfer under limited target domain data.

and autoregressive forecasting. These are expected to help the model learn robust temporal patterns, fault-related features, and domain-invariant representations across heterogeneous PHM datasets.

3) Downstream adaptation under limited target-domain data: After pre-training, the learned backbone representations will be adapted to downstream PHM tasks under limited target-domain data conditions. Different transfer scenarios will be studied, including cross-condition transfer within the same dataset, cross-machine transfer within the same equipment type, and cross-dataset transfer across different industrial systems. These scenarios represent different domain shift problems and can be used to evaluate the transferability and robustness of the proposed methods. For downstream adaptation, lightweight and parameter-efficient fine-tuning strategies, such as adapters, low-rank adaptation (LoRA), prompt tuning, and prefix tuning, will be explored to reduce the dependence on large labeled target-domain datasets. Domain adaptation techniques, such as feature alignment, adversarial learning, and contrastive adaptation, may also be integrated with the pre-trained backbone to further improve performance in the target domain.

The final objective of this research is to develop a PHM-oriented foundation model framework that supports fault diagnosis, anomaly detection, health assessment, and remaining useful life prediction across different industrial domains. By combining large-scale PHM pre-training or existing pre-trained backbone adaptation, and efficient downstream fine-tuning, this research aims to provide a more transferable and

data-efficient framework for PHM applications under limited target-domain data conditions.

3.2. Prior Work and Preliminary Results

The author has several prior publications related to domain transfer and fault diagnosis in PHM. In the first work, multiple deep learning architectures, including multiscale residual convolutional neural networks (CNNs), hybrid long short-term memory (LSTM) networks, and transformer-based models, were integrated into an ensemble learning framework for gearbox fault diagnosis under varying working conditions (Su & Lee, 2025). Transfer learning was employed to adapt the models to new fault classes and operating conditions without requiring full retraining. Experimental results on the PHM North America 2023 gearbox dataset demonstrated improved robustness and generalization across unseen fault scenarios. In the second work, a partial domain adaptation framework was developed for practical industrial settings where only healthy target-domain samples are available during training. The framework integrated a CNN-based feature extractor with maximum mean discrepancy (MMD) loss to align feature distributions between source and target domains (Su et al., 2025).

More recently, the research has expanded toward pre-trained model-based domain adaptation for gearbox fault diagnosis. Initially, a CNN-DistilBERT adaptation network was developed to investigate whether existing pre-trained transformer models can improve transferable representation learning for industrial time-series data. The framework combined a lightweight convolutional feature extractor with a frozen

DistilBERT encoder and MMD-based domain alignment (Ji et al., 2025). Building upon this work, the research was further extended to investigate different pre-trained transformer backbones, including DistilBERT, GPT-2, and T5, for cross-domain fault diagnosis under limited target-domain data conditions. In addition, different domain alignment strategies, including multiple kernel maximum mean discrepancy (MK-MMD) and triplet loss, were developed to further improve domain-invariant representation learning. Preliminary experimental results across multiple domain adaptation tasks (same tasks as in (Ji et al., 2025)), as shown in Table 1, indicate that pre-trained transformer-based frameworks can significantly improve cross-domain robustness and fault classification performance, even when only healthy target-domain samples are available during training. These findings suggest that pre-trained transformer architectures provide a promising direction for reliable and data-efficient PHM domain transfer.

Table 1. Average accuracy (%) on each domain adaptation task across different pre-trained backbones.

Task	DistilBERT	GPT-2	T5
T1	96.43	94.59	93.75
T2	96.48	96.27	96.43
T3	87.33	79.52	90.61
T4	97.52	94.57	99.36
T5	90.90	90.74	92.22
T6	99.13	99.27	99.58

3.3. Timeline

The overall research timeline begins in September 2023 and is expected to culminate in a dissertation defense by May 2028.

Phase 1: Initial Research and Domain Adaptation Development (Completed)

Sep, 2023 – Aug, 2024 (12 months): Conducted literature review on transfer learning, domain adaptation, and intelligent fault diagnosis for PHM applications. (Completed)

Sep, 2024 – May, 2025 (9 months): Initial studies focused on transfer learning and ensemble learning frameworks for gearbox fault diagnosis under varying operating conditions. (Completed)

June, 2025 – May, 2026 (9 months): Developed partial-domain adaptation frameworks for industrial fault diagnosis under healthy-only target-domain data. Later, expanded the research toward pre-trained model-based domain adaptation using transformer backbones such as DistilBERT, GPT-2, and T5. (Completed)

Phase 2: Pre-Trained Model and Foundation Model Development

June, 2026 – Dec, 2026 (7 months): Investigate advanced

pre-trained model architectures and efficient adaptation strategies for industrial PHM tasks. Explore different time-series tokenization methods, parameter-efficient fine-tuning techniques, and domain adaptation approaches under limited target-domain data.

Jan 2027 – Dec 2027 (12 months): Develop PHM-oriented foundation model frameworks through large-scale pre-training using heterogeneous industrial datasets. Investigate self-supervised learning strategies and evaluate transferability across different industrial domains, machines, and operating conditions for downstream tasks.

Phase 3: Dissertation Preparation, and Defense

Jan, 2028 – Mar, 2028 (3 months): Prepare journal manuscripts, finalize dissertation chapters, and start dissertation writing.

Apr, 2028 – May, 2028 (3 months): Complete dissertation writing, committee review, revisions, and dissertation defense.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Education through the Fund for the Improvement of Postsecondary Education (FIPSE) under Grant No. P116S230014.

REFERENCES

- 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.
- Ji, D.-Y., Su, H., Tsuruta, S., Arimizu, D., Hachiya, Y., Wakimoto, K., & Lee, J. (2025). Multi-class gearbox fault diagnosis via pre-trained model-based domain adaptation with healthy-only target data. In *Phm society asia-pacific conference* (Vol. 5).
- Shi, Y., Ying, X., & Yang, J. (2022). Deep unsupervised domain adaptation with time series sensor data: A survey. *Sensors*, 22(15), 5507.
- Su, H., Ji, D.-Y., Tsuruta, S., Arimizu, D., Hachiya, Y., Wakimoto, K., & Lee, J. (2025). Partial domain adaptation for intelligent machinery fault diagnosis: Leveraging healthy-only target data for multi-class classification. In *Annual conference of the phm society* (Vol. 17).
- 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).
- Wang, C., Wang, Z., Liu, Q., Dong, H., Liu, W., & Liu, X. (2025). A comprehensive survey on domain adaptation for intelligent fault diagnosis. *Knowledge-Based Systems*, 114109.
- 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).
- 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.
- Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., . . . others (2025). A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. *International Journal of Machine Learning and Cybernetics*, 16(12), 9851–9915.