Integrating Few-Shot Learning and Pre-trained Models into Similarity-Based PHM using Small Data in Complex Engineering Systems

Takanobu Minami[1,2], Lee Jay[1]

[1]*Center for Industrial Artificial Intelligence, Department of Mechanical Engineering, University of Maryland, College Park, MD, USA*

minamitu@umd.edu, leejay@umd.edu

[2]*Komatsu Ltd., Tokyo, Japan*

ABSTRACT

Prognostics and Health Management (PHM) is vital for complex engineering systems, yet its data-driven solutions are often hampered by the "small data problem"—a scarcity of labeled fault data in industrial settings. This limitation restricts the training and generalization of machine learning models and is compounded by varying operational conditions that reduce the relevance of historical data and pre-trained models. This research introduces a research framework to tackle these small data challenges in PHM. The primary objective is to develop a robust and adaptable PHM methodology by enhancing and synergistically integrating similarity-based Few-Shot Learning (FSL) with large-scale pre-trained time-series models. The research will focus on two main thrusts. First, it aims to improve the generalization capabilities of FSL frameworks by addressing limitations such as noise robustness, domain shift adaptability, and generalization to novel faults across diverse PHM domains. This involves developing noise-robust feature extraction, integrating domain adaptation techniques, and exploring expressive similarity metrics. Second, the study will investigate the effective adaptation of state-of-the-art pre-trained time-series models (e.g., TimesNet) for PHM tasks under data scarcity, focusing on efficient fine-tuning and synergistic integration with the enhanced FSL approaches. The author's prior success in a PHM data challenge using a similarity-based method for spacecraft systems provides preliminary validation. This research is expected to deliver an enhanced PHM framework for high-accuracy diagnostics with limited data, contributing generalized FSL models, systematic methods for leveraging pre-trained models in PHM, and advancing the practical deployment of intelligent PHM solutions.

1. INTRODUCTION

Prognostics and Health Management (PHM) plays a pivotal role in ensuring the reliability and operational efficiency of complex engineering systems. While recent advancements in machine learning, particularly deep learning, have significantly enhanced PHM capabilities, a persistent challenge in real-world industrial applications is the small data problem. This issue arises from the infrequent occurrence of fault events and the high costs associated with collecting and labeling comprehensive health monitoring data. Consequently, training robust and generalizable data-driven models becomes exceedingly difficult, hindering the development of widely applicable diagnostic and prognostic frameworks. Furthermore, varying operational conditions exacerbate this problem by rendering historical data and pre-trained models less effective.

This research aims to address these critical challenges in PHM by proposing a novel framework centered on two primary pillars: (1) enhancing the generalization capabilities of similarity-based Few-Shot Learning (FSL) approaches, and (2) effectively adapting and integrating large-scale pre-trained time-series models for PHM tasks in small data environments. By synergistically combining these strategies, this work seeks to enable high-accuracy and reliable PHM systems even when labeled data is scarce.

2. RELATED WORK AND CHALLENGES

Various approaches have been explored to tackle the small data problem in PHM, including data augmentation, transfer learning, domain adaptation (Zhou et al., 2020), and FSL. FSL, in particular, has garnered significant attention due to its focus on learning to generalize efficiently from very few labeled samples. Similarity-based metric learning and meta-learning techniques, such as Prototypical Networks (Snell et al., 2017), Matching Networks (Vinyals et al., 2016), Relation Networks (Sung et al., 2018), and Model-Agnostic Meta-Learning (MAML) (Finn et al., 2017), have shown promise by effectively capturing inter-class relationships in the feature space.

However, a significant portion of existing FSL research has been evaluated on limited benchmark datasets, often focusing on specific domains like bearing fault diagnosis (Liang et al., 2023). Consequently, the generalization capability of these methods to diverse fault modes, systems with different physical characteristics, or real-world operational environments with inherent variabilities remains insufficiently understood. Adapting to domain shifts and entirely new, unseen fault types presents a major hurdle for practical deployment.

Concurrently, inspired by successes in natural language processing (Brown et al., 2020), pre-trained foundation models, especially those tailored for time-series data (e.g., TimesNet by Wu et al., 2023, MOMENT, Chronos), are emerging as powerful tools. These models, trained on massive and diverse datasets, can learn complex temporal dependencies and patterns, offering significant potential for PHM tasks such as fault diagnosis and Remaining Useful Life (RUL) prediction. Nevertheless, research on effectively adapting these large pre-trained models to specific, often data-scarce, PHM domains is still in its nascent stages. Developing efficient fine-tuning strategies, incorporating domain-specific knowledge, ensuring interpretability, and exploring synergistic combinations with FSL are critical research avenues.

This background underscores the need for an integrated approach that transcends the limitations of individual methods, combining the strengths of FSL in data-efficient learning with the powerful representations learned by pre-trained models.

3. PROPOSED METHODOLOGY

This research proposes a comprehensive framework to address the challenges of PHM in small-data environments, structured around two primary research objectives, each broken down into specific tasks.

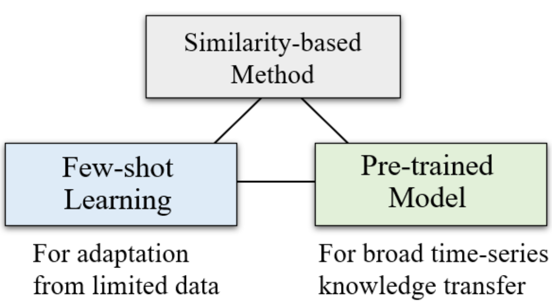


Figure 1. Core Components of Proposed Methodology

3.1. Enhancing the Generalization of Similarity-Based Few-Shot Learning

This objective begins with a systematic evaluation of existing similarity-based FSL frameworks (e.g., Siamese Networks, Prototypical Networks) beyond conventional bearing datasets, using diverse public benchmark datasets representing various engineering systems (e.g., rotating machinery, hydraulic systems, aerospace propulsion systems) with distinct physical phenomena and fault modes. This evaluation will pinpoint key challenges FSL models face, specifically: (a) robustness to various types of noise prevalent in real-world environments, (b) adaptability to domain shifts arising from different operational conditions or equipment variations, and (c) generalization to novel, unseen fault types not present in the training data.

To address these multifaceted challenges, this research will develop and evaluate several enhanced methodologies. Firstly, noise-robust feature extraction techniques will be pursued. This involves incorporating sophisticated feature engineering approaches alongside deep learning architectures, such as those leveraging attention mechanisms or employing robust loss functions, specifically designed to mitigate the impact of noise inherent in real-world data and to extract more salient and discriminative fault features. Secondly, the integration of domain adaptation techniques will be a key focus. This aims to explicitly embed domain adaptation concepts within the FSL frameworks. Strategies will include employing adversarial learning to cultivate domain-invariant feature representations, thereby reducing the gap between training and target domains, and regularizing the learning process by penalizing discrepancies between source and target domain distributions. Finally, the research will explore expressive similarity metrics and hybrid approaches. This moves beyond conventional similarity measures like Euclidean or cosine distances to investigate advanced metrics, such as learnable Mahalanobis distances or kernel methods, which are better suited to capturing complex local data structures and non-linear relationships. Furthermore, hybrid models that synergistically combine the strengths of different FSL paradigms, for instance, by integrating metric learning with prototype-based methods, will be investigated to achieve superior performance and generalization.

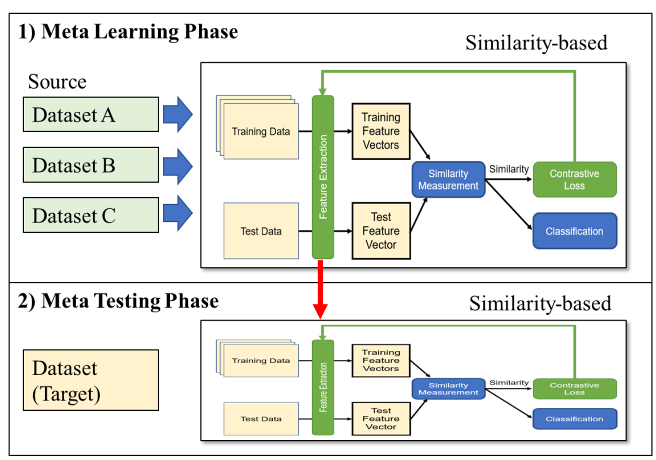


Figure 2. An Example of Conceptual Framework for Similarity-Based Few-Shot Learning

3.2. Adaptation and Integration of Pre-trained Time-Series Models

This objective focuses on applying state-of-the-art pre-trained time-series models (e.g., TimesNet, MOMENT, Chronos) to diverse PHM fault diagnosis tasks to assess their baseline performance and identify adaptation challenges in small-data settings (e.g., domain mismatch, risk of overfitting during fine-tuning, computational cost). To effectively leverage these powerful models for PHM and explore synergies with FSL, the following strategies will be pursued:

Efficient Transfer Learning and Fine-tuning Strategies: Developing and applying parameter-efficient fine-tuning techniques (e.g., adapter modules, LoRA) and progressive unfreezing strategies to mitigate overfitting on small target datasets while efficiently transferring knowledge.

Synergistic Integration with FSL: Developing integrated approaches that leverage the complementary strengths of pre-trained models and FSL. This includes using pre-trained models as powerful backbone feature extractors for FSL frameworks or applying FSL meta-learning strategies (i.e., "learning to learn") to the task adaptation process of pre-trained models, aiming for high diagnostic accuracy even with extremely limited data.

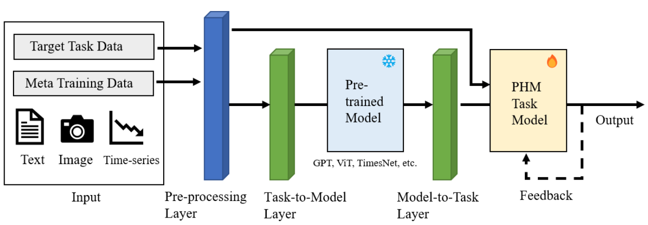


Figure 3. Conceptual Structure of the Similarity-based Method with Pre-trained Model Framework

3.3. Integrated Framework Validation

The elemental techniques developed under objectives 3.1 and 3.2 will be combined to construct an integrated PHM framework optimized for small-data scenarios. The efficacy of this framework will be rigorously validated using multiple public benchmark datasets (e.g., CWRU bearing data, SECOM semiconductor manufacturing data, NASA turbofan engine degradation data) through a comprehensive evaluation protocol. Metrics will include standard N-shot K-way classification accuracy for FSL, robustness under noisy conditions, and cross-domain generalization performance.

4. EXPECTED CONTRIBUTIONS AND IMPACT

This research is poised to make several significant contributions to Prognostics and Health Management (PHM), particularly for prevalent small-data challenges. Primarily, it will deliver an enhanced fault diagnosis framework, a robust, integrated system synergizing advanced Few-Shot Learning (FSL) with adapted pre-trained time-series models for high performance in data-scarce scenarios.

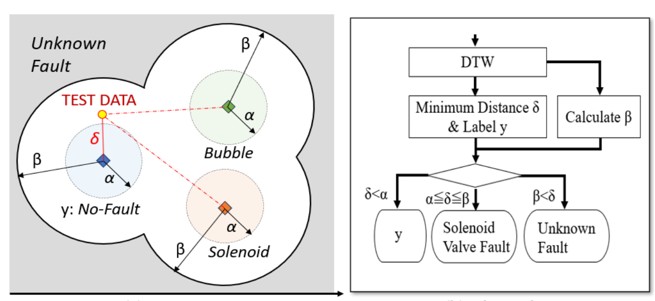
Furthermore, the work will advance generalized similarity-based FSL models, developing frameworks with improved adaptability across diverse industrial domains. This includes enhancing capabilities to address domain shifts and detect novel, unseen fault types. Another key outcome will be efficient adaptation methods for pre-trained models, providing systematic methodologies to tailor large-scale time-series models for specific industrial fault diagnosis, thus facilitating their use when domain-specific data is limited.

Additionally, this study will offer an elucidation of synergies between FSL and pre-trained models, providing insights into the benefits of combining these paradigms and identifying optimal integration strategies for maximizing diagnostic accuracy and data efficiency. Finally, a comprehensive evaluation and benchmarking on diverse public datasets will provide clear, reproducible performance metrics, establishing new baselines for small-data PHM and offering a valuable resource to the research community.

Collectively, these contributions are anticipated to significantly expand the reliability and applicability of PHM systems in real-world industrial settings. By enabling more accurate diagnostics with limited data, this research will contribute to more efficient maintenance, reduced downtime, and more sustainable industrial operations.

5. PRIOR WORK AND PRELIMINARY RESULTS

The author has prior experience directly relevant to the proposed research, particularly in developing and applying similarity-based methods for PHM. A key example is our work for the PHM Asia-Pacific 2023 Conference Data Challenge, which focused on the health assessment of spacecraft propulsion systems. For this challenge, we developed a two-stage methodology. The first stage employed a similarity-based model, leveraging Dynamic Time Warping (DTW) on an identified information-rich initial segment (regime) of the time-series pressure data, to classify the system's health state. The second stage utilized a model incorporating physics-inspired features to further diagnose solenoid valve faults.



(a) Concept (b) Flow Chart

Figure 4. Similarity-Based Classification Method Proposed in PHM AP Data Challenge 2023

This approach demonstrated high classification accuracy during cross-validation with the training dataset and, critically, performed effectively on the unseen test dataset. This test set included data from a new spacecraft not present in the training set, as well as unknown fault types. Our methodology secured a top-tier rank in the competition, underscoring the robustness and adaptability of the similarity-based classification combined with domain-informed feature engineering, especially in scenarios with limited and complex data. These preliminary results strongly support the foundational premise of this proposal: that similarity-based approaches offer a promising avenue for developing data-efficient and generalizable PHM solutions. This prior success, achieved by the author, provides a valuable starting point, practical insights, and validated techniques for the current research proposal, particularly for enhancing FSL generalization.

6. TIMELINE

The proposed research is planned intensively to culminate in a dissertation defense by July 2026. The timeline commences from May 2025.

Phase 1: Focused Development and Initial Integration

May,'25 – July (3 mo.): Develop and validate core enhanced FSL methodologies (noise robustness, domain adaptation); adapt pre-trained models for key PHM tasks.

Aug - Oct (3 mo.): Prototype the integrated PHM framework; conduct preliminary experiments on core benchmarks.

Nov - Dec (2 mo.): Refine framework based on prototype testing; draft initial dissertation sections/conference paper.

Phase 2: Comprehensive Validation and Manuscript Preparation

Jan, '26 - Mar (3 mo.): Conduct rigorous validation of the refined framework on diverse benchmarks; perform comparative analysis (robustness, generalization).

Apr - May (2 mo.): Finalize experiments and analysis; draft and submit a significant journal manuscript; continue dissertation writing.

Phase 3: Dissertation Completion and Defense

June (1 mo.): Final Dissertation Writing and Committee Review.

July,'26 (1 mo.): Revisions and Dissertation Defense.

7. CONCLUSION

This research proposal presents an integrated approach to tackle the critical small-data problem in PHM for complex engineering systems. By focusing on enhancing the generalization of similarity-based FSL and effectively adapting and integrating pre-trained time-series models, the proposed framework aims to significantly improve the adaptability and generalization capabilities of diagnostic models. The author's prior success in leveraging similarity-based methods for a complex PHM challenge provides strong confidence in this research direction. The successful execution of this research, following the outlined timeline, is expected to advance the practical deployment of PHM technologies across diverse industrial domains, even when faced with limited data availability.

REFERENCES

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901.

Finn, C., Abbeel, P., & Levine, S. (2017, July). Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning* (pp. 1126-1135). PMLR.

Liang, X., Zhang, M., Feng, G., Wang, D., Xu, Y., & Gu, F. (2023). Few-shot learning approaches for fault diagnosis using vibration data: a comprehensive review. *Sustainability*, *15*(20), 14975.

Minami, T., & Lee, J. (2023, September). Phm for spacecraft propulsion systems: Similarity-based model and physics-inspired features. In *Phm society asia-pacific conference* (Vol. 4, No. 1).

Snell, J., Swersky, K., & Zemel, R. (2017). Prototypical networks for few-shot learning. *Advances in neural information processing systems*, *30*.

Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P. H., & Hospedales, T. M. (2018). Learning to compare: Relation network for few-shot learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1199-1208).

Vinyals, O., Blundell, C., Lillicrap, T., & Wierstra, D. (2016). Matching networks for one-shot learning. *Advances in neural information processing systems*, *29*.

Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., & Long, M. (2022). Timesnet: Temporal 2d-variation modeling for general time series analysis. *arXiv preprint arXiv:2210.02186*.

Zhou, K., Yang, Y., Hospedales, T., & Xiang, T. (2020). Learning to generate novel domains for domain generalization. In *Computer vision–ECCV 2020: 16th European conference, glasgow, UK, August 23–28, 2020, proceedings, part XVI 16* (pp. 561-578). Springer International Publishing.