A Multidisciplinary Framework for Vibration-Based Gear Fault Diagnosis Using Experiments, Modeling, and Machine Learning

Lior Bachar[1], Jacob Bortman[1]

[1]*Departmant of Mechanical Engineering, Ben-Gurion University of the Negev, Beer-Sheva, 8410501, Israel*

liorbac@post.bgu.ac.il
jacbort@bgu.ac.il

ABSTRACT

Vibration-based gear diagnosis is crucial for ensuring the reliability of rotating machinery, making the monitoring of gear health essential for preventing costly downtime and optimizing performance. This study proposes a multidisciplinary framework to enhance fault diagnosis, that aligns with digital twin principles by integrating experiments, dynamic modeling, physical preprocessing, and machine learning. Within this framework, we focus on three core procedures: domain adaptation to reduce discrepancies between measured and simulated data; physical preprocessing, grounded in in-depth investigations dictating signal processing and feature engineering techniques; and learning algorithms, encompassing the process of training AI-based models. The framework is benchmarked through a comprehensive case study of localized tooth fault diagnosis, using controlled-degradation tests and realistic simulations. First, we detect faults using unsupervised learning algorithms; then, we use zero-shot-learning for classifying between localized and distributed faults; finally, we adopt a few-shot-learning strategy for severity estimation. Above all, this hybrid framework aligns with the accelerating field of physics-informed machine learning, by combining physical knowledge and advanced algorithmics with machine learning. This contributes to the PHM community by offering valuable insights into integrating different aspects of research, thereby enhancing performance in diagnosis tasks.

1. INTRODUCTION

Vibration-based gear diagnosis has made significant strides over the years, resulting in a concise general framework that typically encompasses data collection, signal processing, feature extraction, and health indicator construction (Kumar, Gandhi, Zhou, Kumar, and Xiang, 2020; Kundu, Darpe, and Kulkarni, 2021). However, challenges persist, particularly in the accelerating field of digital twins and physics-informed machine learning (DENG et al., 2023). The availability of labeled faulty measured data remains limited, a shortcoming that must be acknowledged in any data-driven diagnosis strategy. Furthermore, each fault type manifests differently, necessitating tailored methods to overcome these differences and recognize their unique characteristics in the signature in early stages. Condition-based maintenance, as illustrated in Figure 1, typically encompasses diagnosis, including fault detection, classification, and severity estimation; and prognosis, which involves remaining useful life estimation (Kumar et al., 2020). This work contributes a comprehensive framework for gear fault diagnosis that aligns with the growing area of digital twins. Dynamic models can be instrumental, both for bridging theoretical insights with practical applications (Mohammed & Rantatalo, 2020) and for generating synthetic training data (Bachar et al., 2023). While other approaches do not necessarily rely on dynamic modeling, we choose to incorporate them as indispensable assets (Dadon, Koren, Klein, and Bortman ,2018). By combining physical preprocessing with synthetic data, the proposed framework aims to improve generalizability of AI-based algorithms for fault diagnosis. Section 2 outlines the proposed framework. Section 3 summarizes extensive controlled-degradation tests in gears. Section 4 introduces the general flow of incorporating dynamic models. Section 5 presents a case study on localized tooth fault diagnosis, benchmarking the effectiveness of the proposed framework.

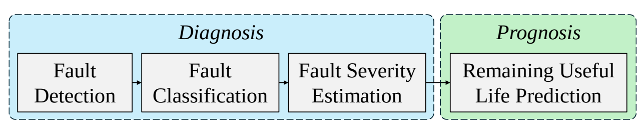

Figure 1. General stages of PHM.

2. THE MULTIDISCIPLINARY FRAMEWORK

Figure 2 presents a block diagram of the proposed framework, which assumes that users have abundant labeled healthy data, limited labeled faulty data (if any), and a rich database of both healthy and faulty simulated data. The first step involves using domain adaptation to enhance the simulated data, aiming to minimize discrepancies between simulated (source) and measured (target) data. Next, the vibration data undergoes meticulous physical preprocessing using signal processing and feature engineering techniques. Finally, learning algorithms are trained to make predictions for the desired diagnosis task. The following subsections provide detailed insights into each core stage.

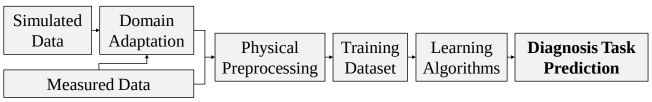

Figure 2.Block diagram of the multidisciplinary framework.

2.1. Domain Adaptation

Figure 3 illustrates the domain adaptation strategy, which specifically aims to reduce discrepancies between simulated and measured data. The raw simulated signal passes through estimated transfer functions (TFs) derived from the target machine, emulating the transmission path between excitation and sensor. The TF is approximated by assuming a minimum phase (mp) for the cepstrum of the background spectrum (Oppenheim and Schafer, 1999), as illustrated in Figure 4 and Equation 1 (Bachar et al., 2023).

$$TF(f) = \exp(DFT(cepstrum_{mp}(background(f))) \qquad (1)$$

This approach enables the creation of synthetic data with feature trends that closely resemble those of the actual machine, even when faulty data of the target is unavailable, as showcased in Section 5. With the abundance of healthy measured data, a set of transfer functions (TFs) can be estimated to enhance and augment simulations. For example, a single simulation of a pitted gear can be enhanced to generate dozens of realistic signals across different machines. While this method provides a reasonable approximation, we acknowledge its limitations in more complex cases or under varying operational conditions, and therefore use it primarily to demonstrate domain adaptation within our framework. More broadly, this strategy lays the foundation for digital twinning in future endeavors.

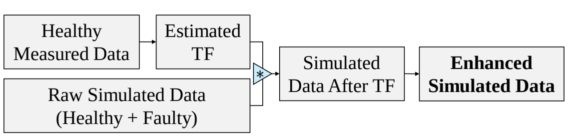

Figure 3. Block diagram of the domain adaptation.

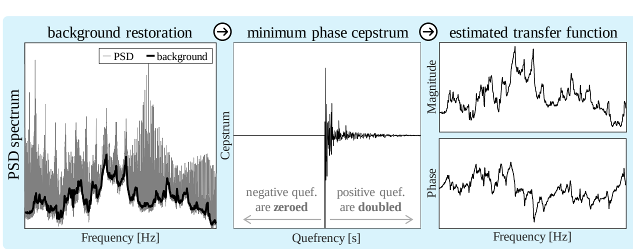

Figure 4. Illustration of the transfer function estimation.

2.2. Physical Preprocessing

Figure 5 illustrates the physical preprocessing strategy, which leverages physical knowledge to enhance the performance of predictive algorithms. Gear vibrations exhibit characteristic patterns, thoroughly investigated by experts over the years. Gear signal processing relies on understanding their dynamic behavior, with synchronous averaging (sa) as a key algorithm (Bechhoefer & Kingsley, 2009; Kumar et al., 2020; Sharma & Parey, 2016). The sa signal captures a shaft's full cycle, isolated from non-synchronized components and noise, offering a strong basis for feature extraction. The difference (diff) signal highlights modulation phenomena linked to localized faults such as pitting, derived by filtering out gearmesh harmonics and close sidebands from the sa (Bachar et al., 2021, 2022), as illustrated in Figure 6. Conversely, the harmonics signal, which includes only gearmesh harmonics, is informative mainly for monitoring distributed wear faults (Randall, 1982). We extracted features from the sa domain, and categorize them as follows:

- *Shape-based features*: Emphasize sharp impulsive responses from faults, including sa kurtosis, diff kurtosis, sa envelope skewness, and diff envelope skewness.
- *Energy-based features*: Highlight energy variations associated with faults, including sa rms, sa envelope rms, diff rms, diff envelope rms, and spectral energy at modulation sidebands far from gearmesh harmonics.

Meticulous feature engineering further enhances predictive power through techniques such as scaling, dimensionality reduction, feature selection, aggregation, etc.

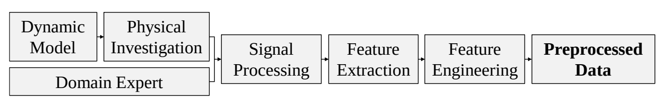

Figure 5. Block diagram of the physical preprocessing.

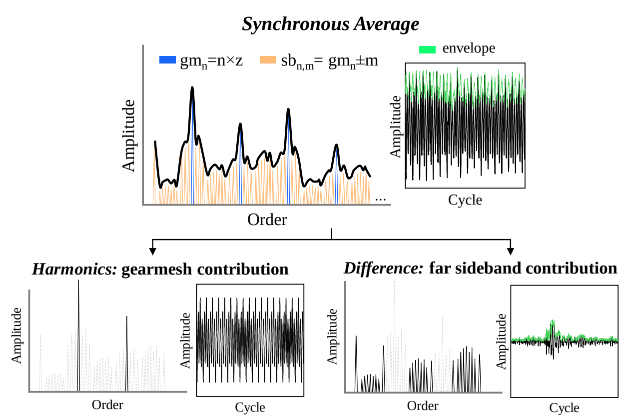

Figure 6. Scheme of synchronous average analysis, accompanied by a signal of a pitted gear.

2.3. Learning Algorithms

Figure 7 illustrates the learning algorithm strategy, aligning with the growing field of physics-informed machine learning, which integrates physical knowledge with AI to enhance accuracy, generalization, and efficiency in fault diagnosis of complex systems (DENG et al., 2023; Lei et al., 2020).

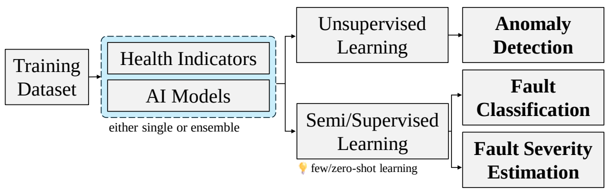

Figure 7. Block diagram of the learning algorithms.

AI models and/or health indicators are trained using preprocessed vibration data (Wang, Tsui, and Miao, 2017) to perform the desired diagnosis task (see Figure 1):

1. *Anomaly Detection*: This unsupervised learning task focuses on distinguishing between healthy and faulty samples. Common algorithms include local outlier factor (Yuan et al., 2023; Zhang et al., 2024), and isolation forest (Wang, Li, Liu, and Yang., 2022). Early detection can be challenging, requiring detectors to extract sensitive features (Bachar et al., 2023). Therefore, training relies solely on data from the target, without integrating data from other sources, to better understand the healthy population manifold.

2. *Fault Classification*: Gear faults are generally categorized into localized (e.g., pitting) and distributed (e.g., wear) (Eugene E. Shipley, 1967). This task involves pinpointing the defective gear and determining the fault type. With coprime numbers of teeth on each wheel, locating the faulty wheel in the sa is straightforward. However, classifying the fault type is challenging due to the limited availability of faulty data, particularly for incipient faults. Supervised learning approaches, such as few-shot (Liang et al., 2023; Wang et al., 2020) and zero-shot learning (Pourpanah et al., 2022; Zhang et al., 2022), can potentially be applied, utilizing synthetic data to capture patterns associated with each fault type.

3. *Fault Severity Estimation*: This task typically relies on supervised learning methods, such as kNN and Random Forest, for severity classification or regression. The scarcity of faulty data poses challenges for zero-shot learning (Cerrada et al., 2018), so we adopt a few-shot learning approach (Orozco & Roberts, 2020; Wang et al., 2020), leveraging limited labeled faulty data from the target machine to normalize synthetic data. Domain adaptation and physical preprocessing help the estimator handle minor discrepancies with the target machine for accurate prediction. Additionally, training separate estimators for each fault type improves performance and allows for precise severity definitions, such as worn area for wear or tip loss for tooth breakage.

3. CONTROLLED-DEGRADATION TESTS

Vibration data can be collected through two test types: endurance tests, which measure data continuously over time as faults evolve naturally, and controlled-degradation tests, which induce faults artificially and measure data over severity, without necessarily having prior knowledge about degradation mechanisms (Dadon, Koren, Klein, and Bortman, 2019; Feng, Ji, Ni, and Beer, 2023; Kumar et al., 2020). Investigating the degradation rate over time in endurance tests may be critical for prognosis, but the lack of ground truth information about the system's health affects their performance in diagnosis. The absolute control over fault morphology in controlled-degradation tests is crucial for validating dynamic models and diagnostic algorithms (Liang, Zuo, and Feng, 2018). However, repetitive assembly operations during fault seeding introduce minor structural variations that are not related to its health. The proposed framework integrates data measured through controlled-degradation tests with simulated data, prioritizing control over severity rather than degradation over time. Representative tests on spur gears by Dadon et al. (2019) and Bachar et al. (2022, 2024) are shown in Figure 8.

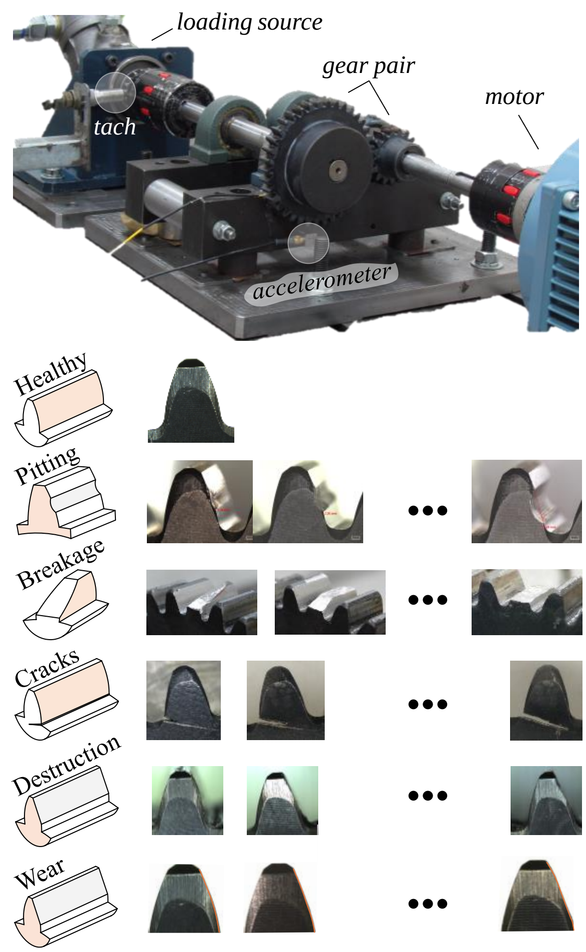

Figure 8. A typical gear test apparatus (top) and various controlled-degradation tests conducted (bottom).

4. DYNAMIC MODELS FOR GEARS

The major strength of the proposed framework is its strong reliance on simulated data. We utilize a dynamic model for gear vibrations, first introduced by Dadon et al. (2018). The model accommodates any combination of gear pair, operational conditions, manufacturing errors, and health status. Vibration signals are derived from the solution to the equations of motion, accounting for the time-variant gearmesh stiffness. Figure 9 presents a general flow of gear dynamic modeling. Experimental validation involves visual inspection of the vibration signal and qualitative comparison of feature trends across fault severity, demonstrating similar behavior between simulation and experiments, as shown in Figure 10 for the test detailed in Section 5 and Figure 11. The utilized model has been experimentally validated for healthy gear, and for various faults including pitting, breakage, destruction, and wear (Dadon et al., 2018, 2019; Bachar, et al., 2024). A potential concern with incorporating dynamic models is how to find a suitable model and validate it when labeled faulty data is rare. While this is a valid limitation, we assume that the typical behavior of gear faults remains consistent regardless of the scale of the test setup. With proper domain adaptation, dynamic models validated on small-scale setups may generate synthetic data that closely resemble unseen faulty data from the target machine. Additionally, we encourage the sharing of organized application programming interfaces (APIs) for existing dynamic models to make them publicly accessible, thereby facilitating the implementation of our proposed framework.

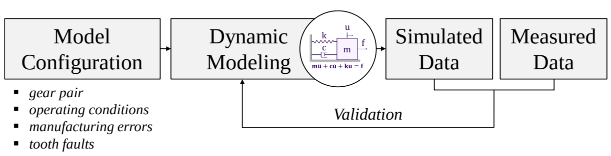

Figure 9. Block diagram of dynamic modeling and experimental validation.

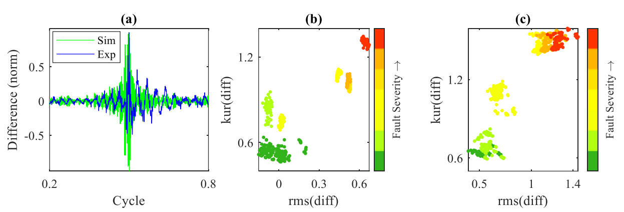

Figure 10. Validation of tooth destruction faults: (a) Comparison of normalized difference signals for the most severe fault; logarithmic maps of kur(diff) vs. rms(diff) across fault severity for (b) experiments and (c) simulations.

5. CASE STUDY – MONITORING LOCALIZED TOOTH DESTRUCTION FAULTS IN SPUR GEARS

The proposed framework is demonstrated through a case study on localized tooth destruction faults, with severity labels defined by the damaged area (mm²). Figure 12 illustrates this study flow, covering all diagnosis stages from Figure 1: fault detection, classification, and severity estimation.

5.1. Experimental & Simulated Datasets

The test setup in this study comprises a spur gearbox with a driving pinion (z_p=18T) and driven gear (z_g=35T), both with a DIN8 finish. The input speed is 45rps, and the output load is 10Nm. Localized and distributed faults are introduced into the driven gear. Data is sampled at 50kS/s, covering z_p cycles of the driven shaft. Datasets are categorized as follows:

- *Healthy measured data*: Contains data from the same healthy gear, with variance introduced through repetitive assembly operations. It is used for domain adaptation and training algorithms. The induced variance prevents overly optimistic performance from similarity between train and test sets, known as training-test leakage.
- *Faulty measured data*: Contains a wide range of fault severities (localized and distributed) used for evaluation and few-shot learning, accounting for the lack of labeled faulty data in real-world scenarios.
- *Simulated data*: A large database of simulated data, enhanced through domain adaptation. It covers a much wider range of fault sizes than the experiment (see Figure 11), utilized to train few/zero-shot learning algorithms.

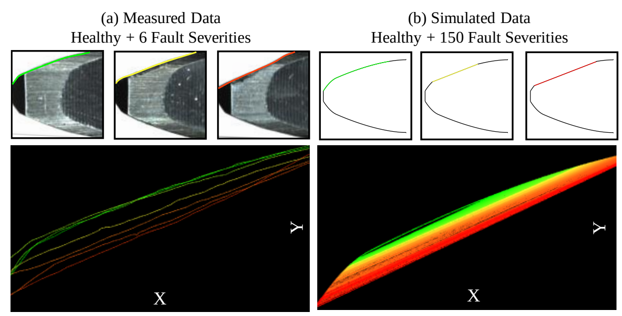

Figure 11. Tooth contours across severity of measured and simulated datasets in the case study.

5.2. Fault Detection

We extract common features based on statistical moments of the vibration sa signal, along with sensitive features from the spectrum, to train an unsupervised isolation forest (iForest). This detector splits the data into trees and assigns an anomaly score based on the average path length needed to isolate each point (Liu, Ting, and Zhou, 2008). The iForest is trained on healthy data and validated using leave-2-out (L2O) cross-validation across assembly operations. Figure 13 presents the anomaly scores (in units of the standard deviation σ of the training scores) plotted against fault size, with error bars included. The results demonstrate the effectiveness of the proposed iForest in accurately detecting faults. Healthy samples, both from the training and test sets, are clustered with low values. While the healthy test samples differ slightly from the training set, likely due to structural differences, these variations (~2σ) are not significant enough to classify them as novel. In contrast, faults are clearly separated, with anomaly scores of at least ~8σ, even for incipient damage.

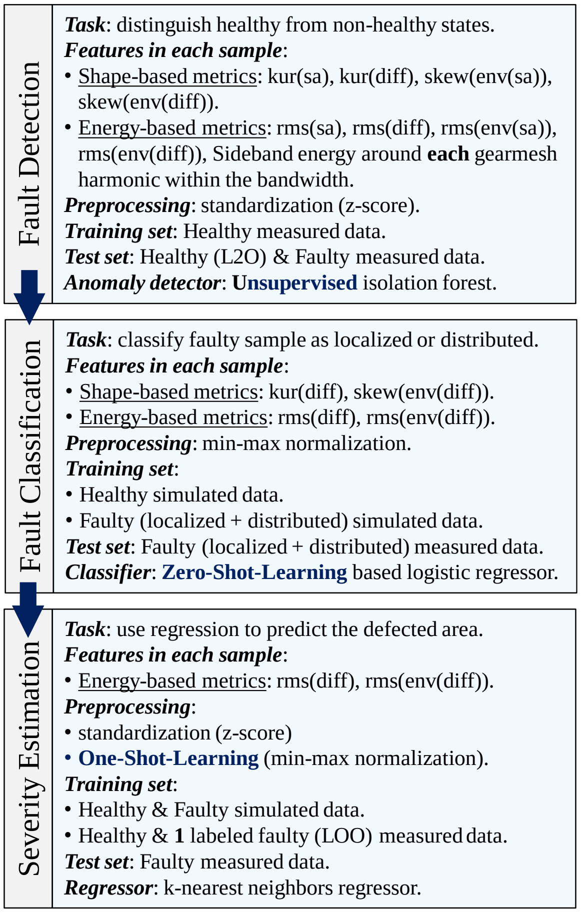

Figure 12. Schematic flow of the case study.

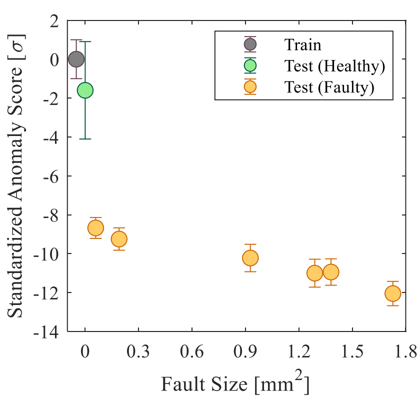

Figure 13. iForest anomaly scores vs. fault size.

5.3. Fault Classification

This stage involves categorizing the fault type for samples identified as anomalies. We train a zero-shot learning classifier to distinguish between localized (y=0) and distributed (y=1) faults, focusing on broader categories rather than specific subtypes. To enable zero-shot learning, we leverage synthetic data that closely resemble the dynamic behavior, such as sharp impulses for localized faults. This approach enables accurate fault classification without the need for training on measured faulty data. We train a logistic regressor (Cox, 1972), using four commonly shape-based and energy-based features used for monitoring localized faults (see Figure 12). Figure 14 illustrates the performance of our zero-shot learning classifier by displaying the predicted probability against each input feature, highlighting their individual correlations with the prediction. The results show that the separation between localized and distributed faults is more pronounced in shape-based features (kurtosis and skewness), likely contributing more to the prediction, as expected. The zero-shot classifier relies on enhanced simulated training data, but performance may decline if there are significant discrepancies or mismatches in operational conditions between the simulated and measured target data. While further investigation is recommended, it is important to note that the classifier broadly distinguishes between localized and distributed faults using a coarse set of features. As a result, it is expected to be less sensitive to minor discrepancies, unlike the detector in the previous stage.

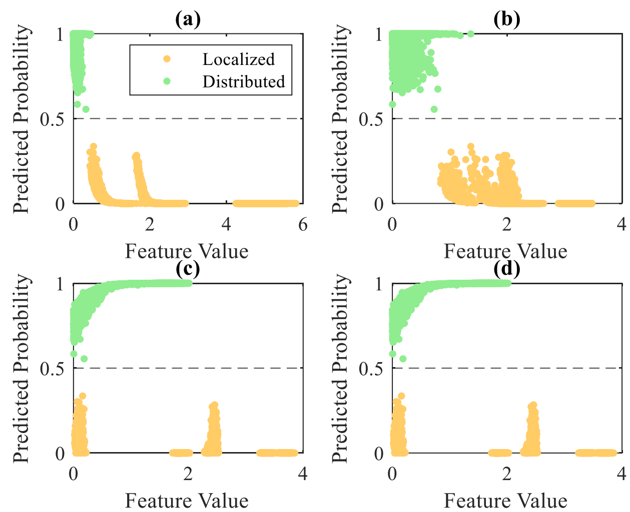

Figure 14. Zero-shot fault classification. Normalized feature value vs. predicted probability: (a) kur(diff); (b) skew(env(diff)); (c) rms(diff); (d) rms(env(diff)).

5.4. Fault Severity Estimation

The severity label is estimated for samples classified as localized using a few-shot learning strategy, extracting two energy-based features: rms(diff) and rms(env(diff)). The training measured data includes healthy (mean $\overline{X}_{exp,h}$) and faulty (mean $\overline{X}_{exp,1fault}$) samples from one severity label only. A kNN regressor is trained on simulated data (X_{sim}), normalized using healthy (mean $\overline{X}_{sim,h}$) and faulty (mean $\overline{X}_{sim,1fault}$) data from the same severity label as in the training measured data, and tested on measured data (X_{exp}), normalized with the training measured data (Equations 2-3).

$$X_{train} = (X_{sim}-\overline{X}_{sim,h})/ (\overline{X}_{sim,1fault} - \overline{X}_{sim,h}) \quad (2)$$

$$X_{test} = (X_{exp}-\overline{X}_{exp,h})/ (\overline{X}_{exp,1fault} - \overline{X}_{exp,h}) \quad (3)$$

Using leave-one-out (LOO) cross-validation, we examine how the single severity label y (mm²) in the training measured data, used for normalization, impacts performance. Figure 15.a shows the mean absolute error (MAE) of the test set against the k-neighbors hyperparameter for different severity labels y (see Figure 11.a). Results suggest that fault severity estimation is more effective in few-shot learning when larger severity labels are used in the training measured data for normalization, regardless of the value of k. This may be due to challenges in generating realistic synthetic data for small faults. Additionally, recall that while scaling simulated and measured data by healthy and same-severity samples allows knowledge transfer, it does not enforce a linear relationship between them, and this limitation may be more pronounced when small severity labels are used for scaling. However, further investigation is needed to confirm these insights. Figure 15.b shows the predicted vs. true labels for the three largest severity labels y in experiment, demonstrating the solid performance in estimating severity with minimal labeled faulty measured data.

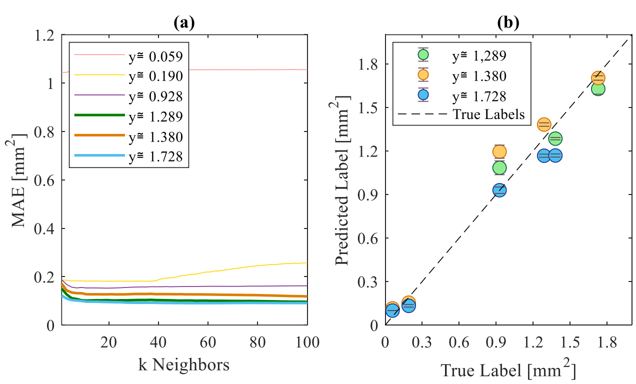

Figure 15. kNN-based severity estimator: (a) MAE vs. k-neighbors for different train severity labels; (b) Predicted vs. true labels for the three largest severity labels in training.

5.5. Benchmarking

The core objective of our framework is to utilize simulated data that closely resembles measured data to train diagnostic algorithms, addressing the shortage of labeled faulty measured data. The premise is that fully supervised training with abundant labeled data would yield ideal performance compared to synthetic data. To benchmark our framework, we specifically assess how fault classification and severity estimation perform in a fully supervised setting with rich labeled faulty measured training data, aiming to demonstrate that our framework achieves comparable results even with minimal or no labeled faulty measured training data.

- *Fault classification* – Figure 16.a presents an rms(diff)-kur(diff) map (normalized by healthy status), comparing measured data of localized and distributed faults across health degradation, with darker markers indicating greater severity. The results indicate a clear, nearly linear separation between fault types, with a more pronounced distinction in kurtosis, as expected. Classic machine learning classifiers like logistic regression, SVM, kNN, or decision trees would perform optimally with this level of separability. Despite the noisier separability of our fault classifier (see Figure 14), we still achieved comparable performance using a zero-shot learning approach trained solely on synthetic data.

- *Fault severity estimation* – Figure 16.b shows the trend of rms(diff) in the measured data across severity labels, showing a clear, nearly linear correlation with health degradation. Classic machine learning regressors, such as linear regression, SVR, kNN, or decision trees, would likely perform accurately if trained on a rich set of labeled faulty measured data that strongly correlates with severity, as illustrated in the figure. Despite the proposed few-shot learning estimator's limitations, such as sensitivity to the severity label in the training data (see Figure 15), we still achieved performance comparable to the benchmarking training dataset, demonstrating the potential of using enhanced synthetic training data for severity estimation.

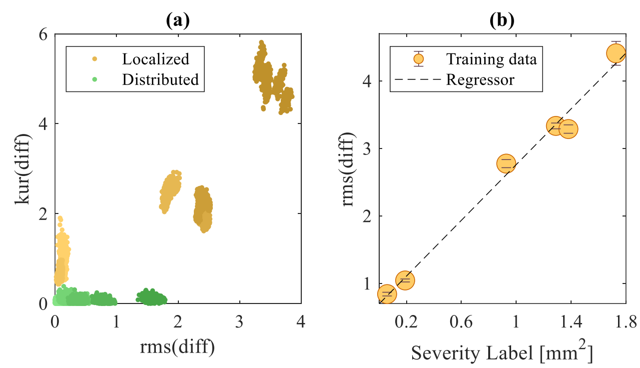

Figure 16. Benchmarking against fully supervised training: (a) fault classification shown by clear separation in normalized rms(diff)-kur(diff) maps; (b) fault severity estimation illustrated by the rms(diff) vs. severity label.

6. CONCLUSION

This study introduces a comprehensive framework integrating experimentation, dynamic modeling, physical preprocessing, and data-driven machine learning techniques for vibration-based gear fault diagnosis. The proposed multidisciplinary framework begins with domain adaptation to minimize discrepancies between simulated and measured data, followed by rigorous preprocessing using advanced signal processing and feature engineering techniques to prepare the data for AI-based algorithms. Through a detailed case study of localized tooth fault diagnosis, we successfully demonstrate the framework's efficacy step-by-step across unsupervised algorithms for anomaly detection, zero-shot-learning algorithms for fault classification, and one-shot-learning algorithms for fault severity estimation. The framework's reliance on simulated data is both its strength and limitation, which can be mitigated by sharing organized APIs and making validated dynamic models publicly available. A qualitative benchmarking demonstrated that the proposed framework effectively overcomes practical challenges, achieving performance comparable to an ideal fully supervised training dataset. By bridging traditional physical methods with realistic synthetic data and classic AI-based approaches, our hybrid methodology enhances gear diagnosis capabilities, and potentially other critical components, such as bearings. This research advances the PHM field and highlights the value of multidisciplinary approaches in optimizing performance and generalizability, particularly in the growing domains of digital twins and physics-informed machine learning.

ACKNOWLEDGEMENT

We gratefully acknowledge the invaluable support of the Pearlstone Foundation.

REFERENCES

Bachar, L., Dadon, I., Klein, R., & Bortman, J. (2021). The effects of the operating conditions and tooth fault on gear vibration signature. *Mechanical Systems and Signal Processing*, *154*, 107508. https://doi.org/10.1016/J.YMSSP.2020.107508

Bachar, L., Klein, R., Tur, M., & Bortman, J. (2022). Fault diagnosis of gear transmissions via optic Fiber Bragg Grating strain sensors. *Mechanical Systems and Signal Processing*, *169*. https://doi.org/10.1016/j.ymssp.2021.108629

Bachar, L., Matania, O., Cohen, R., Klein, R., Lipsett, M. G., & Bortman, J. (2023). A novel hybrid physical AI-based strategy for fault severity estimation in spur gears with zero-shot learning. *Mechanical Systems and Signal Processing*, *204*. https://doi.org/10.1016/j.ymssp.2023.110748

Cerrada, M., Sánchez, R., Li, C., … F. P.-… S. and S., & 2018, undefined. (2018). A review on data-driven fault severity assessment in rolling bearings. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0888327017303242?casa_token=MeMxmyHKbHwAAAAA:TbKPBWg7SxdDjxdbOkhQ2Hv5l_cdwAlw7ZlOvCRBqv6wXH6JNVwfiY__iqGXRTtY5KvuQD-Qwz0

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *34*(2), 187–202. https://doi.org/10.1111/J.2517-6161.1972.TB00899.X

Dadon, I., Koren, N., Klein, R., & Bortman, J. (2018). A realistic dynamic model for gear fault diagnosis. *Engineering Failure Analysis*, *84*. https://doi.org/10.1016/j.engfailanal.2017.10.012

Dadon, I., Koren, N., Klein, R., & Bortman, J. (2019). A step toward fault type and severity characterization in spur gears. *Journal of Mechanical Design*, *141*(8). https://doi.org/10.1115/1.4043367

DENG, W., NGUYEN, K. T. P., MEDJAHER, K., GOGU, C., & MORIO, J. (2023). Physics-informed machine learning in prognostics and health management: State of the art and challenges. *Applied Mathematical Modelling*, *124*, 325–352. https://doi.org/10.1016/J.APM.2023.07.011

Eugene E. Shipley. (1967). *GEAR FAILURES*. Penton Publishing Co. https://www.xtek.com/wp-content/uploads/2018/05/xtek-gear-failures.pdf

Feng, K., Ji, J. C., Ni, Q., & Beer, M. (2023). A review of vibration-based gear wear monitoring and prediction techniques. In *Mechanical Systems and Signal Processing* (Vol. 182). https://doi.org/10.1016/j.ymssp.2022.109605

Kumar, A., Gandhi, C. P., Zhou, Y., Kumar, R., & Xiang, J. (2020). Latest developments in gear defect diagnosis and prognosis: A review. In *Measurement: Journal of the International Measurement Confederation* (Vol. 158). https://doi.org/10.1016/j.measurement.2020.107735

Kundu, P., Darpe, A. K., & Kulkarni, M. S. (2021). A review on diagnostic and prognostic approaches for gears. *Structural Health Monitoring*, *20*(5), 2853–2893. https://doi.org/10.1177/1475921720972926/ASSET/IMAGES/LARGE/10.1177_1475921720972926-FIG7.JPEG

Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. In *Mechanical Systems and Signal Processing* (Vol. 138). https://doi.org/10.1016/j.ymssp.2019.106587

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. https://doi.org/10.3390/su152014975

Liang, X., Zuo, M. J., & Feng, Z. (2018). Dynamic modeling of gearbox faults: A review. *Mechanical Systems and Signal Processing*, *98*, 852–876. https://doi.org/10.1016/J.YMSSP.2017.05.024

Liu, F., Ting, K., international, Z. Z.-2008 eighth ieee, & 2008, undefined. (2008). Isolation forest. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/4781136/?casa_token=cAMoWKo1ZnkAAAAA:3lFnB_-my6ecFZ3RA8fSioWjkE2kqQrdHzsZueH-F9CJlvHbAGXYl8L132L0k279b5Z7xrTfw6Wo

Mohammed, O. D., & Rantatalo, M. (2020). Gear fault models and dynamics-based modelling for gear fault detection – A review. *Engineering Failure Analysis*, *117*, 104798. https://doi.org/10.1016/J.ENGFAILANAL.2020.104798

Oppenheim, A., Schafer, R., & Buck, J. (1999). *Discrete-time signal processing*. https://dl.acm.org/doi/abs/10.5555/294797

Orozco, B. P., & Roberts, S. J. (2020). Zero-shot and few-shot time series forecasting with ordinal regression recurrent neural networks. *ESANN 2020 - Proceedings, 28th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 503–508.

Pourpanah, F., Abdar, M., Luo, Y., … X. Z.-I. transactions on, & 2022, undefined. (2022). A review of generalized zero-shot learning methods. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/9832795/

Randall, R. B. (1982). A New Method of Modeling Gear Faults. *Journal of Mechanical Design*, *104*(2), 259–267. https://doi.org/10.1115/1.3256334

Wang, D., Tsui, K. L., & Miao, Q. (2017). Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators. *IEEE Access*, *6*, 665–676. https://doi.org/10.1109/ACCESS.2017.2774261

Wang, H., Li, Q., Liu, Y., Machines, S. Y.-, & 2022, undefined. (2022). Anomaly data detection of rolling element bearings vibration signal based on parameter optimization isolation forest. *Mdpi.Com*. https://www.mdpi.com/2075-1702/10/6/459

Wang, Y., Yao, Q., Kwok, J., (csur), L. N.-A. computing surveys, & 2020, undefined. (2020). Generalizing from a few examples: A survey on few-shot learning. *Dl.Acm.Org*, *53*(3). https://doi.org/10.1145/3386252

Yuan, D., Han, Y., Zhu, Y., -, al, Zhang, H., Tang, Q., Xu, D., Liu, S., Ye, Z., Yi, L., Wang, H., Yang, S., Liu, Y., & Li, Q. (2023). A novel abnormal data detection method based on dynamic adaptive local outlier factor for the vibration signals of rotating parts. *Iopscience.Iop.Org*, *18*(11), 11. https://doi.org/10.1088/1361-6501/accbda

Zhang Examiner, Y., Bi, X., & Kedziora, D. (2024). *Development of vibration signal preprocessing and outlier application*. https://lutpub.lut.fi/handle/10024/167465

Zhang, T., Chen, J., Li, F., Zhang, K., Lv, H., He, S., transactions, E. X.-I., & 2022, undefined. (2022). Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0019057821001257?casa_token=rKLklNCfTHYAAAAA:5XaeIkh5IeWd7rNgelPjyC3KfVAwVwwJyUpOLKwSSu3DgG_PnhkEiDJlSar7X-Buko95Xb8Kyus

BIOGRAPHIES

Lior Bachar is currently a current Ph.D. student in mechanical engineering in BGU-PHM LAB at Ben-Gurion University of the Negev, Beer-Sheva, Israel. His primary academic interest is vibration-based diagnosis of gears. His research, published in numerous papers, employs hybrid physical data-driven based approaches, integrating experimentations, signal processing, dynamic modeling, and AI applications. Lior completed with honors his bachelor's degree and master's degree in mechanical engineering in Ben-Gurion University of the Negev.

Jacob Bortman is currently a full Professor in the department of mechanical engineering and the head of the PHM Lab in Ben-Gurion University of the Negev. Retired from the Israeli air force as brigadier general after 30 years of service with the last position of the head of material directorate. Chairman and member of several boards: director of business development of Odysight Ltd, Chairman of the board of directors, Selfly Ltd., board member of Augmentum Ltd., board member of Harel finance holdings Ltd., Chairman of the board of directors, Ilumigyn Ltd. Editorial board member of: “Journal of Mechanical Science and Technology Advances (Springer, Quarterly issue)”. Head of the Israeli organization for PHM, IACMM - Israel Association for Comp. Methods in Mechanics, ISIG - Israel Structural Integrity Group, ESIS - European Structural Integrity Society. Received the Israel National Defense prize for leading with IAI strategic development program, Outstanding lecturer in BGU, The Israeli prime minister national prize for excellency and quality in the public service - First place in Israel. Over 80 refereed articles in scientific journals and in international conference