Gear Diagnostics Based On Transfer Learning Methodologies and Digital Twinning

Henrique Duarte Vieira de Sousa[1, 2] and Konstantinos Gryllias [1, 2]

[1] *Department of Mechanical Engineering, KU Leuven*
henriqueduarte.vieiradesousa@kuleuven.be
konstantinos.gryllias@kuleuven.be

[2] *Flanders Make @ KU Leuven*
Celestijnenlaan 300, Box 2420, 3001 Leuven, Belgium

ABSTRACT

In 2023, renewable energy sources represented 24.5% of total energy consumption within the EU (Eurostat, n.d.-b). This figure reflects progress towards the target set by the European Wind Power Action Plan, which aims for 42.5% of the EU's energy consumption to be derived from renewable sources by 2030 (Eurostat, n.d.-a), with wind energy representing one of the main contributors to this transition. However, data compiled in (Santelo, de Oliveira, Maciel, & de Almeida Monteiro, 2022) indicate that wind turbines (WTs) experience an average of 0.402 failures per year, with each failure resulting, on average, in 130 hours of downtime per turbine. Gearbox failures, in particular, represent one of the most frequent and impactful failure modes, significantly contributing to the overall levelized cost of energy (LCOE).

As a result, the implementation of condition-based monitoring (CM) for high-risk components is of considerable importance to maximize operational availability and minimize downtime. CM methodologies focus on three core tasks: (1) fault detection, (2) diagnosis, and (3) prognosis. These are generally pursued through two main approaches: signal processing techniques and machine learning (ML) algorithms.

Focusing on the latter approach, the advent of deep learning (DL), together with the increasing volume of data generated by modern systems, has brought a significant capability to infer complex relationships within the data. This enables high accuracy across all three components of condition-based monitoring (CM) requiring less extensive domain expertise than what is often necessary for signal processing techniques. However, as pointed out in (Li et al., 2022), the application of DL in CM is hindered by two main challenges: the scarcity of high-quality labeled data and the shift in distribution between training and testing datasets, both of which limit the performance of DL-based models.

To mitigate these limitations, transfer learning (TL) has been introduced. As described by (Pan & Yang, 2010), TL allows for differences between training and testing domains and tasks. A domain is defined by a feature space and a marginal probability distribution, $D = \{\mathcal{X}, P(X)\}$, where $X = \{x_1, \ldots, x_n\} \in \mathcal{X}$. A task is defined by a label space and a predictive function, $\mathcal{T} = \{\mathcal{Y}, P(y|x)\}$. In contrast to standard machine learning formulations — where training and testing must share the same domain and task — TL permits differences in feature mappings, marginal distributions, label spaces, and objectives. In TL, three categories are established, depending on the characteristics of the domains and tasks, according to (Pan & Yang, 2010). They are: (1) inductive; (2) transductive; and (3) unsupervised.

Within the context of CM, transductive TL is the predominant paradigm for classification tasks, since labeled data in the source domain is essential for identifying the fault and its location, while the central challenge lies in addressing the distributional shift between the source and target domains. Out of the possible methodologies, domain adaptation remains the most useful, assuming a shared feature space between domains and aiming to identify a common latent representation in which the marginal distributions are aligned.

However, domain adaptation has advanced beyond the classical formulation—known as closed-set domain adaptation—which assumes identical label spaces and focuses solely on domain shift. As datasets have grown in complexity, this assumption has become increasingly restrictive. To address this, the notion of category gap has been introduced, as discussed in (Farahani, Voghoei, Rasheed, & Arabnia, 2021), where the label space $\mathcal{Y}$ is permitted to differ across domains. This generalization, illustrated in Figure 1, enables the use of more various datasets, thereby enhancing model robustness and flexibility.

Moreover, the integration of data from multiple source

domains—referred to as multi-source domain adaptation—further extends the applicability of TL in CM. Despite the promise of these more general approaches, current research, as summarized in (Li et al., 2022), indicates that closed-set, single-source domain adaptation remains the most commonly adopted strategy in the field.

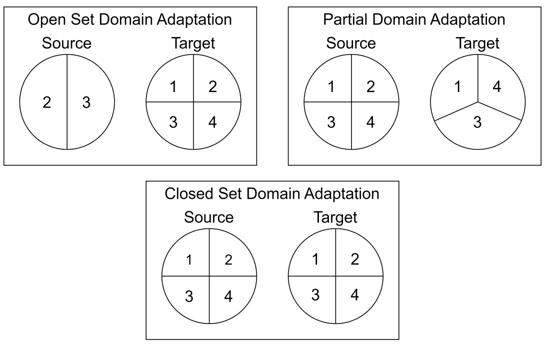


Figure 1. Types of domain adaptation according to (Farahani et al., 2021).

An alternative approach within the TL field is the integration of digital twins (DTs) to generate synthetic healthy and faulty data. In this framework, a virtual representation of the physical system serves as the source domain. Three primary modeling approaches can be employed in ascending order of both fidelity and computational cost: (1) phenomenological / analytical models; (2) multibody dynamic models; or (3) high-fidelity finite element models (FEM). The use of the more complex solutions offer a higher degree of realism, with the ability of simulating the transmission path of fault-induced vibrations from their origin to sensors mounted on the housing. In contrast, phenomenological and analytical models are computationally more efficient, allowing for rapid generation of large synthetic datasets, albeit at the cost of physical accuracy. Given the focus of this research on CM and ML application, the current state-of-the-art (SOTA) in digital twin technologies is sufficient to create a detailed dataset that can simulate the behavior of the real system under healthy and faulty conditions.

Beyond modeling techniques, sensor diversity remains an interesting characteristic to explore in CM. While accelerometers and encoders dominate the field, alternative sensing technologies, such as fiber optic sensors, have untapped potential. These have already been adopted in structural health monitoring but remain in a infant stage for CM applications. Notably, fiber Bragg grating (FBG) sensors present a promising avenue for further investigation. Lastly, how to combine all the different type of information for both DTs and ML algorithms is a topic that can be delved into.

The present research aims to develop and validate a methodology for fault classification within gearboxes, with a particular focus on gear-related anomalies. Validation will be conducted using both measurements acquired in in-house dedicated test rigs and publicly available datasets relevant to WTs. The anticipated contributions to the field of prognostics and health management (PHM) include:

1. The development of a robust end-to-end framework for fault classification, from raw sensor input to final diagnostic output, with computational efficiency suitable for real-time application;
2. A seamless integration between the DT of the gearbox, the simulation of synthetic faults, and the generation of TL-compatible datasets;
3. A hybrid diagnostic method that fuses domain-specific signal processing features with deep learning representations;
4. The incorporation of heterogeneous sensor data — primarily from accelerometers, encoders, and fiber optic sensors (specifically, FBG sensors) — to enhance diagnostic robustness.

The overall framework of the proposed methodology can be illustrated using a DT of an in-house gearbox with an example developed by (Liu & Gryllias, 2022). During the course of this research, a new method will be developed, and multiple types of digital twins (DTs) will be evaluated, effectively replacing the workflow shown in Figure 2.

The achievement of the research objectives will follow the preliminary structured plan outlined below:

1. **Literature review.** This task will be carried out throughout the entire research timeline and focuses on three main areas: (a) current SOTA in TL and ML methods; (b) gearbox modeling, with special attention to gear-level fault modeling; and (c) signal processing techniques and health indicators (HIs).
2. **Transfer learning.** The author will study various SOTA methodologies as described in (Li et al., 2022) and beyond, to understand the current vanguard within CM, and build upon them. Within the topic of TL, the author wants to create a flexible methodology that is able to generalize to more domains than the one representing the real asset of the DT. The research will explore open-set domain adaptation, recognizing that real-world systems rarely follow the assumption that source and target label spaces are identical and its performance will be benchmarked. Multi-source domain adaptation will also be examined, combining real and synthetic domains to enhance knowledge transfer. In addition, the effects of having labeled or unlabeled data in the target domain will be explored to maximize robustness. The author will study positive transfer learning to determine the conditions under which source domains contribute constructively to performance, and will investigate methods to

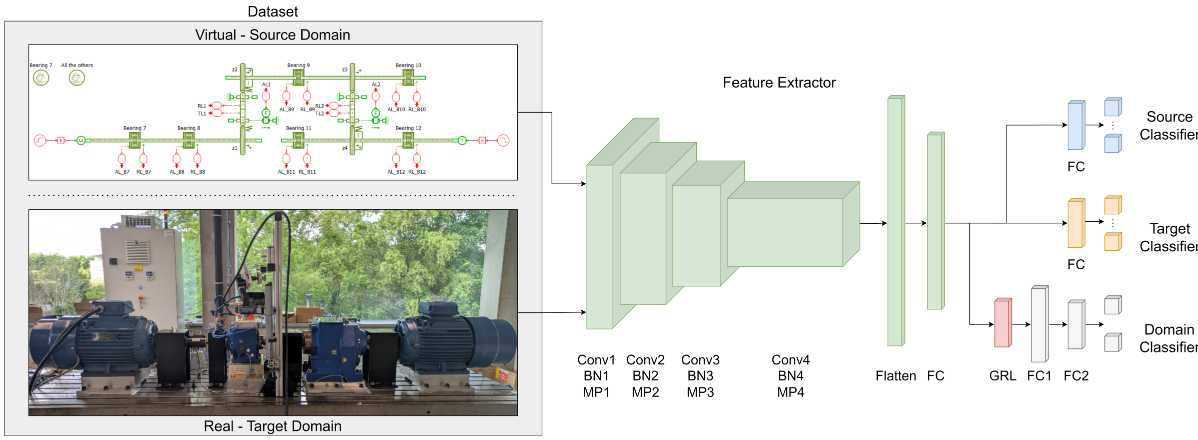


Figure 2. An example for the methodology using AMESim and the methodogy proposed by (Liu & Gryllias, 2022).

control dataset imbalance, which may otherwise distort classifier behavior. In order to avoid establishing the ML algorithm as just a black box, the author will also dive into the realm of explainable AI to allow for more transparency between the model's decisions and the framework's user. Lastly, a detailed investigation into the ML components, such as network architecture, optimization algorithms, training dynamics, will be performed to ensure accurate and reliable decision-making.

3. **Signal processing study.** This task involves reviewing SOTA signal processing techniques across time, frequency, and time-frequency domains. Signal preparation methods such as segmentation, normalization, and domain-specific preprocessing like Time Synchronous Averaging will be studied first. Subsequent efforts will focus on: (a) characterizing how faults manifest in time and frequency signals to inform feature extraction; (b) evaluating statistical indicators like RMS, variance, kurtosis, and skewness to enrich model inputs; and (c) applying more advanced tools such as cyclostationary analysis (Antoni, 2007), envelope spectrum analysis, wavelet transforms, and short-time Fourier transforms for robust feature extraction. The goal is to provide a robust signal processing backbone for the ML framework.

4. **Modeling study.** The author will develop DTs of gearboxes, first using data from experimental test rigs and later extending the models to wind turbines, with the knowledge available from the SOTA. Furthermore, common gear and bearing faults, such as pitting, tooth crack, and spalls, will be studied and simulated within the DT, providing a fully synthetic dataset that will train the algorithm into correctly classifying the data coming from the real asset. The resulting data will reflect sensor measurements from accelerometers, encoders, and fiber optic sensors by modeling the transfer path from component to housing (with this step not being possible for the phenomenological and analytical models). Lastly, the author will look into model updating techniques to fine-tune the behavior of the model, as data arrives from the real system. The modeling aspect of the research will be performed at various levels, starting from phenomenological models all the way to multi-body software and finite element solutions.

5. **Validation.** This task will occur throughout the research process. Accelerated life testing will be conducted to induce realistic gear and bearing faults under various operational conditions. Data from diverse sensor types will be used to validate the proposed ML framework, both in training and in testing phases.

The test rig that will be used is presented in Figure 3. This setup will not only support experimental validation of the proposed methodology but will also provide the necessary data for training the machine learning algorithm.

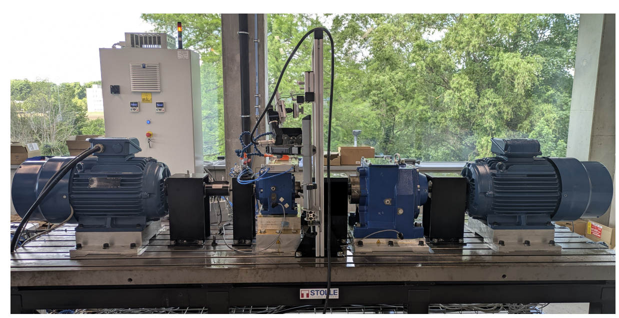


Figure 3. Back-to-Back Gearbox Test Rig.

The test rig was designed with the primary objective of accelerating the development of gear defects while enabling visual observation of their onset. Beyond this, it provides valuable diagnostic information, including the measurement of transmission error (TE), which will contribute directly to the objectives of this research.

The rig consists of two identical induction motors, each capable of delivering 30 kW at a nominal speed of 3000 rpm. One motor acts as the driver, rotating the drivetrain, while the other provides the load. This configuration is enabled by the presence of two gearboxes in the system: the test gearbox and a secondary ”drive” gearbox that functions as a load multiplier. This arrangement not only permits the use of smaller motors to achieve target load levels but also enables both speed and torque control using identical motors, provided that both gearboxes share similar gear ratios.

Both gearboxes are two-stage helical parallel-shaft configurations. The test gearbox has a gear ratio of 5.12 given by $z_1 = 48,\ z_2 = 57,\ z_3 = 13,\ z_4 = 56$, while the drive gearbox has a gear ratio of 5.27 defined by $z_1 = 37,\ z_2 = 44,\ z_3 = 14,\ z_4 = 62$. The test gearbox is lubricated using a mineral oil, with temperature monitored at both the inlet and outlet, whereas the drive gearbox employs synthetic oil. Both gearboxes utilize traditional oil bath lubrication systems.

The primary fault mode of interest is pitting, with additional focus on spalls and tooth cracks. These faults will be artificially accelerated by overloading the system based on the design formulas established in (American Gear Manufacturers Association, 2004), targeting the most vulnerable component, i.e., the second-stage pinion.

The test rig is equipped with four encoders: one on each motor, and one at both the input and output shafts of the test gearbox. The motor encoder on the speed motor supports the closed-loop control of speed, while the gearbox encoders enable the global calculation of TE for the gearbox. To support visual diagnostics, the test gearbox has been modified to allow optical access to the second-stage pinion. This enables visual inspection during acquisition time and the creation of a visual dataset synchronized with sensor data, allowing precise identification of fault initiation and evolution.

Additional instrumentation includes a torque sensor positioned between the test and drive gearboxes, which contributes to torque loop control. Multiple accelerometers are distributed across the rig: on the motors, the base plate, the drive gearbox, and most importantly, on the test gearbox. Two tri-axial accelerometers are mounted on the test gearbox - one on the input side and the other on the output side — to capture high-resolution vibration data critical for fault diagnosis.

An in-house software has been developed to support automated testing with the various data acquisition systems. This software controls the speed and load profiles (both stationary and time-varying), interfaces with the traditional data acquisition system and the camera, and automates the collection of vibration and visual data. With an initial offset between data types, vibration data is recorded every 10 minutes, while visual data is acquired hourly. Synchronization is achieved by temporarily reducing the rotational speed during image capture to ensure clarity. Continuous monitoring of the output-side accelerometer, a key sensor, is also implemented as a safety measure. After acquisition, all real-time data is streamed to an online monitoring platform where SOTA and in-house algorithms are applied for advanced condition monitoring.

In summary, this paper outlines the motivation for the research, reviewed the relevant SOTA in TL and CM, and identified some current research gaps. Moreover a dedicated test rig that will be used for methodological development and experimental validation has been described in detail. Finally, a structured research plan has been proposed, with the ultimate objective of developing a robust and scalable methodology combining ML and DTs for fault diagnostics of WT gearboxes, thereby contributing meaningfully to the field of PHM.

ACKNOWLEDGMENT

The authors acknowledge the financial support of the Clean Energy Transition Partnership, funded from the European Union’s Horizon Europe research and innovation programme under grant agreement no. GA 101069750, and VLAIO, via the Hybrid Wind project.

REFERENCES

American Gear Manufacturers Association. (2004). *Fundamental rating factors and calculation methods for involute spur and helical gear teeth* [ANSI Standard]. (ANSI/AGMA 2001-D04)

Antoni, J. (2007). Cyclic spectral analysis in practice. *Mechanical Systems and Signal Processing*, *21*(2), 597-630. doi: https://doi.org/10.1016/j.ymssp.2006.08.007

Eurostat. (n.d.-a). *Commission sets out immediate actions to support the european wind power industry.* Retrieved 2023-10-24, from https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5185

Eurostat. (n.d.-b). *Renewables account for 24.5% of EU energy use in 2023.* Retrieved 2024-12-19, from https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20241219-3

Farahani, A., Voghoei, S., Rasheed, K., & Arabnia, H. R. (2021). A brief review of domain adaptation. In R. Stahlbock, G. M. Weiss, M. Abou-Nasr, C.-Y. Yang, H. R. Arabnia, & L. Deligiannidis (Eds.), *Advances in data science and information engineering* (p. 877-894). Cham: Springer International Publishing. doi: https://doi.org/10.1007/978-3-030-71704-9_65

Li, W., Huang, R., Li, J., Liao, Y., Chen, Z., He, G., ... Gryllias, K. (2022). A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges. *Mechanical Systems and Signal Processing*, *167*, 108487. doi: https://doi.org/10.1016/j.ymssp.2021.108487

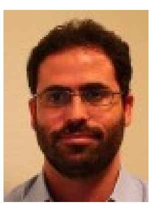
Liu, C., & Gryllias, K. (2022). Simulation-driven domain adaptation for rolling element bearing fault diagnosis. *IEEE Transactions on Industrial Informatics*, *18*(9), 5760-5770. doi: 10.1109/TII.2021.3103412

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345-1359. doi: https://doi.org/10.1109/TKDE.2009.191

Santelo, T. N., de Oliveira, C. M. R., Maciel, C. D., & de Almeida Monteiro, J. R. B. (2022). Wind turbine failures review and trends. *Journal of Control, Automation and Electrical Systems*, *33*(2), 505-521. doi: https://doi.org/10.1007/s40313-021-00789-8

BIOGRAPHIES

Henrique Duarte Vieira de Sousa received his Bachelor and Master of Mechanical Engineering from University of Porto, Portugal in 2022 and 2024, respectively. In 2024, he joined the Condition Monitoring research group in the Department of Mechanical Engineering at KU Leuven, Belgium as a Ph.D. researcher. His research interests lie in the areas of condition monitoring, transfer learning, and digital twins.



Konstantinos Gryllias received the Diploma and Ph.D. degrees in Mechanical Engineering from the National Technical University of Athens, Greece, in 2004 and 2010, respectively. He is currently a Professor of Vibro-Acoustics of Machines and Transportation Systems with the Department of Mechanical Engineering, KU Leuven, Belgium. He also serves as the Manager of the University Core Lab Flanders Make@KU Leuven Motion Products, Belgium. His research interests include condition monitoring, signal processing, prognostics, and health management of mechanical and mechatronic systems.