

Domain Adaptation Digital Twin for Rolling Element Bearing Prognostics

Chenyu Liu^{1,2}, Alexandre Mauricio^{1,2}, Junyu Qi^{1,2}, Dandan Peng^{1,2} and Konstantinos Gryllias^{1,2}

¹ *Department of Mechanical Engineering, Faculty of Engineering Science, KU Leuven*

² *Dynamics of Mechanical and Mechatronic Systems, Flanders Make
Celestijnenlaan 300, BOX 2420, 3001 Leuven, Belgium
konstantinos.gryllias@kuleuven.be*

ABSTRACT

Artificial Intelligence (AI) is escalating in data-driven condition monitoring research. Traditional expert knowledge-based Prognostics and Health Management (PHM) processes can be smartened up with the assistance of various AI techniques, such as deep learning models. On the other hand, current deep learning based prognostics suffers from the data deficit issue, especially considering the varying operating conditions and the degradation modes of the components in practical industrial applications. With the development of simulation techniques, physical-knowledge based digital twin models give engineers access to a large amount of simulation data at a lower cost. These simulation data contain the physical characteristics and the degradation information of the component. In order to accurately predict the Remaining Useful Life (RUL) during the degradation process, in this paper, a bearing digital twin model is constructed based on a phenomenological vibration model. A Domain Adversarial Neural Network (DANN) is used to achieve the domain adaptation target between the simulation and the real data. Regarding the simulation data as the source domain and the real data as the target domain, the DANN model is able to predict the RUL without any priori knowledge of the labelling information. Based on the validation results from real bearing run-to-failure experiments, the proposed method is able to get the minimum RUL prediction error compared against state-of-the-art methods.

1. INTRODUCTION

Predictive Maintenance (PdM) strategy is a popular solution for various industries to achieve cost-efficiency by reducing unnecessary repairs and unplanned downtime. Today, the advancement of Internet-of-Things (IoT) and the data analytics technologies have become the new driving forces behind

the PdM push. Powerhouse companies in the manufacturing and energy sectors are marrying their sensory data with cloud computing platforms, racing to deploy data-driven prognostic tools on mechanical assets.

Artificial Intelligence (AI) is considered as a game-changer for data-driven prognostics. Traditional expert knowledge-based predictive models can be smartened with AI-powered assistance, and upgraded to an advanced prognostic ecosystem with minimal human interference, proactive decision-making, and high prediction accuracy (Lei et al., 2018). In the frame of AI, Deep Learning (DL) has sparked prognostic algorithms' development in light of the robust nonlinear modeling capability. In the last years, a wide variety of DL models have been proposed, targeting an accurate prediction of the Remaining Useful Life (RUL) of critical mechanical components such as bearings and gears, as reviewed by Khan & Yairi (2018).

Despite the heuristic prediction results, state-of-the-art DL prognostics models suffer from the data deficit issue, especially under varying operating conditions and degradation modes for heterogeneous machinery fleets in the real industry. One solution is to conduct large-scale, fleet-wise run-to-failure experiments to generate the data mountain to improve the generalization ability of the constructed DL model. However, such luxury experiments are not always feasible, considering the critical components' long lifespans in high-valued assets.

Simulation-based Digital Twin (DGT) has great potential to ease the reliance on real-world tests and is gaining fast momentum in the condition monitoring research field. In conjunction with the on-line sensory measurements, DGT models are able to represent the operating scenarios under a variety of parameters (Sobie et al., 2018). In the frame of DL-driven condition monitoring, the primary benefit of DGT is to generate simulated system responses with pre-configured faults or degradation processes, giving access to a large amount of simulation data to train the DL model. The gen-

Chenyu Liu 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.

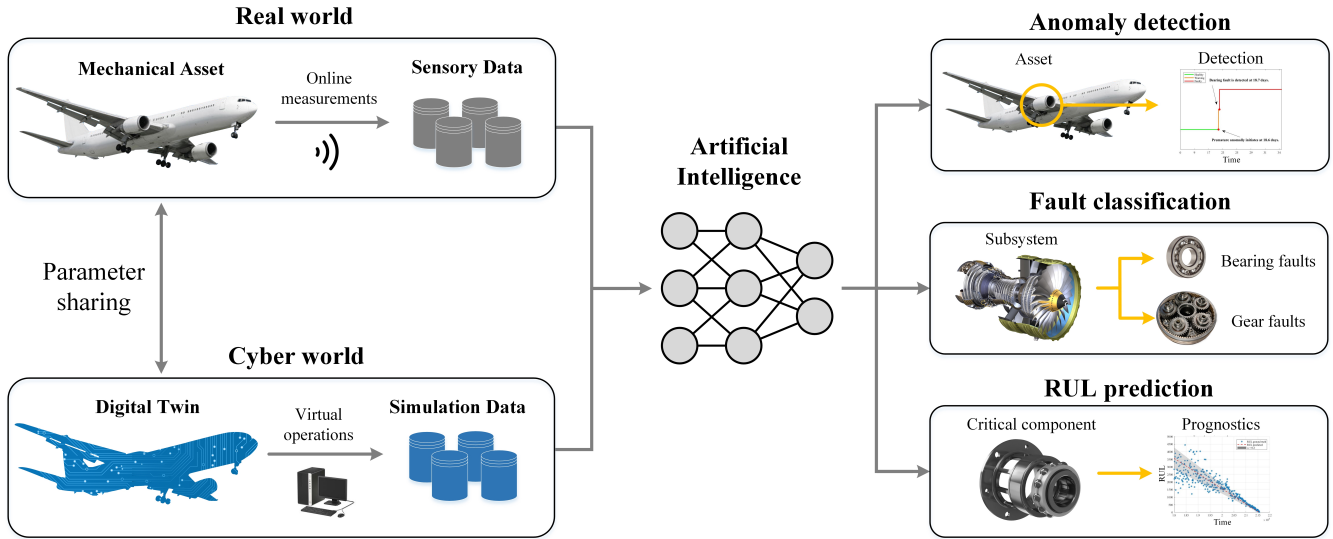


Figure 1. Digital twin-based deep learning for smart maintenance tasks.

eral scheme of DGT-based DL for smart maintenance tasks is depicted in Figure 1.

The pioneering work of the DGT-based AI method was proposed by Gryllias & Antoniadis (2012), where an analytical model was utilized to provide training datasets for a Support Vector Machine (SVM), which was designed for bearing diagnostics. After that, other sophisticated simulation tools were leveraged, and the AI technique was evolving from shallow classifiers to deep models, i.e. deep neural networks. For instance, Sobie et al. (2018) constructed a one-dimensional 3-DOF bearing dynamic model to produce simulated vibration signals for a Convolutional Neural Network (CNN) classifier. Gao et al. (2020) investigated the usage of a Finite Element Model (FEM) to generate faulty bearing features, which was integrated with a Generative Adversarial Network (GAN) to discriminate bearing defects under the hypothesis of insufficient faulty signal information.

These implementations show that the DGT-based DL approach accommodates the requirements in the condition monitoring context. However, the coupling of the DGT and DL models is still at its infancy with limited applications only for classification problems. It has not yet been extended to any prognostic tasks. On the one hand, current DGT models cannot sustain the entire asset degradation processes due to the complex operating environment in practice, which results in low accuracy predictive simulation. On the other hand, the performance of the DL prognostic model is essentially dictated by the distribution mismatch between the simulation data and real measurements, i.e. the domain shift issue.

One solution to tackle this problem is by using the domain adaptation technique to enhance the DL model. Domain Adversarial Neural Network (DANN) is developed as an effective

approach to align the latent feature distributions of the datasets from two different domains (Ganin et al., 2016). Considering the simulation data and real measurements as the source and target domain respectively, DANN adversarially trains a domain discriminator to extract the domain-invariant features, which puts efforts in reducing the domain shift. Recent research from Q. Wang et al. (2019) revealed the DANN method's superiority in bearing fault classification with vibration signals. Chen et al. (2020) also came to a similar conclusion with the CNN-based DANN architecture for end-to-end bearing and gearbox fault diagnostic purposes. For prognostics, da Costa et al. (2020) proposed a DANN model for aero-engine remaining useful life prediction using physical features as inputs.

In this paper, a novel DGT-based DL prognostic model is proposed within the frame of DANN. For the first time, the DANN technique aligns the rolling element bearing physical simulation with run-to-failure experiments for RUL prediction, which carries out the domain adaptation target between the simulation and the real data. The phenomenological vibration model is adopted as the generative DGT, and the Bi-directional Long Short Term Memory (Bi-LSTM) neural network is then employed as the feature extractor in the DANN regime. The results show that the proposed method could get high prediction accuracy in contrast to the non-adapted models.

The rest of the paper is organized as follows: Section 2 introduces the theories of the DGT model and the DANN approach. The proposed method and the experiment are discussed in Section 3 and 4, respectively. The prognostic results are illustrated in Section 5. Some conclusions are summarized in the last section.

2. THEORETICAL PART

2.1. Faulty bearing DGT model

Vibration signals have been proved efficient in reflecting and monitoring the degradation of components, especially for high-dynamic machinery. Thus the proposed DGT model will focus on the simulation of vibration response. The morphologies of rolling element bearing faulty signals are shown in Figure 2, where the faults are mostly observed on the outer race, the inner race, and the rolling elements. By replicating the bearing vibration signals, the DGT model is expected to introduce the critical physical characteristics of the bearing degradation process.

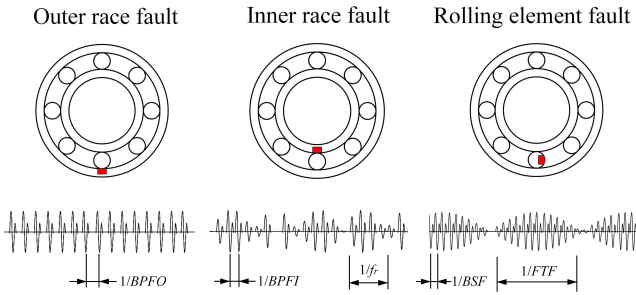


Figure 2. Three types of bearing faults with corresponding time domain vibration responses.

McFadden & Smith (1984) firstly proposed a bearing phenomenological model according to the repetitive impacts and modulations caused by the faults in order to reconstruct the vibration signals. Ho & Randall (2000) improved this model by adding randomness considering the slip phenomenon inside the bearing. Based on the understanding of the quasi-cyclostationarity of bearing signals, Antoni & Randall (2003) proposed a stochastic modelling approach with a more precise simulation of the spectral characteristics. This approach is implemented in this paper as the core DGT model.

For localized defect, the periodic train of impulses could be described as (Antoni, 2007):

$$x(t) = \sum_{i=-\infty}^{+\infty} h(t - iT - \tau_i)q(iT)A_i + n(t) \quad (1)$$

where $h(t)$ represents the impulse response to the measured impact, i is the sequential number of the impact, T is the time interval between two impacts, τ and A are the uncertainties on the inter-arrival time and the magnitude, $q(t)$ and $n(t)$ denote the periodic modulation generated by the load distribution and the background noise, respectively. The randomness of τ and A can be further described with:

$$\begin{aligned} E\{\tau_i \tau_j\} &= \delta_{ij} \sigma_\tau^2, \\ E\{A_i^2\} &= 1 + \delta_{ij} \sigma_A^2 \end{aligned} \quad (2)$$

where δ_{ij} is the Kronecker symbol, σ_τ and σ_A are the stan-

dard deviations. Many studies focusing on the implementation of Equation 1 have demonstrated that the slight random fluctuations of the impulse train could affect the harmonics and result in significant randomness of the signal (Randall et al., 2001; Antoni & Randall, 2003). Consequently, the faulty bearing vibration signal can be described as:

$$x(t) = x_H(t) + x_R(t) \quad (3)$$

where $x_H(t)$ is the weak harmonic component in the low-frequency range, and $x_R(t)$ is the dominating random cyclostationary component in the higher-frequency range. The detailed numerical implementation and the spectral characteristic analysis of the simulated signals can be found at Antoni & Randall (2003).

2.2. Domain Adversarial Neural Network

As one of the domain adaptation techniques, DANN is able to align the simulation efforts with the deep learning model. The fundamental theory of DANN can be traced back to the research from Ganin & Lempitsky (2014) and Ganin et al. (2016). It is designed to encounter the domain shift problem, i.e., the distribution mismatch of training and testing dataset. Inspired by the Generative Adversarial Networks (GANs), DANN uses adversarial training to construct a domain-invariant feature space for both the source and target domain data. These features are simultaneously sent to a domain classifier and to a label predictor. By learning from the source domain data and the associated labels, the goal of DANN is to map a function that can precisely label the target domain data.

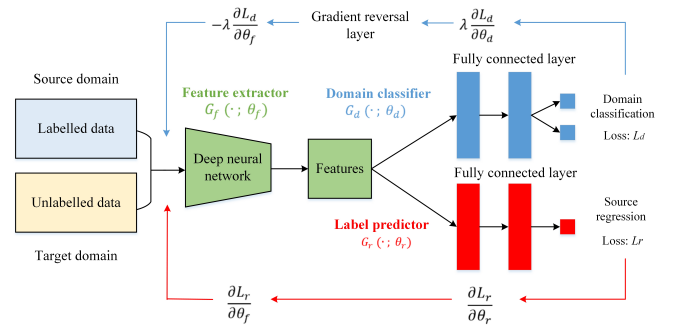


Figure 3. The architecture of DANN for regression.

As shown in Figure 3, the DANN architecture includes three parts: the feature extractor $G_f(\cdot; \theta_f)$, the domain classifier $G_d(\cdot; \theta_d)$ and the label predictor $G_r(\cdot; \theta_r)$, where θ_f , θ_d and θ_r are respectively their hyper-parameters. The G_f is used to create the feature space needed for the following networks based on the mixture of source and target domain samples. In the forward propagation, the features are sent to the G_d to classify whether they come from the source or the target

domain. The loss of the G_d can be described as:

$$L_d(\theta_f, \theta_d) = L_d(G_d(G_f(x; \theta_f); \theta_d)) \quad (4)$$

where x is the input. According to Ganin et al. (2016), the training target of G_d is to reduce the \mathcal{H} -divergence between the two domains, which can be fulfilled using a Gradient Reversal Layer (GRL). During the back propagation process, the GRL is inserted between the domain classifier and the feature extractor to reverse the training targets, i.e., G_f is optimized to extract domain-invariant features from the two domains, but G_d is optimized to discriminate as much as possible their belonging domains. GRL can be simply implemented by multiplying -1 to the gradient without introducing other hyper-parameters. When the classifier can no longer identify the exact domain of a sample, the domain-invariant feature space aligning the two distributions is successfully constructed.

Since the labels are available for the source domain data, the features are simultaneously sent to G_r , which is trained in a supervised way. The loss function of G_r is described as:

$$L_r(\theta_f, \theta_r) = L_r(G_r(G_f(x; \theta_f); \theta_r)) \quad (5)$$

Combining the domain classifier and the label predictor, the total loss of DANN can be defined as follows according to Ganin et al. (2016):

$$L(\theta_f, \theta_d, \theta_r) = \frac{1}{n} \sum_{i=1}^n L_r^i(\theta_f, \theta_r) - \lambda \left(\frac{1}{n} \sum_{i=1}^n L_d^i(\theta_f, \theta_d) + \frac{1}{n'} \sum_{i=1}^{n'} L_d^i(\theta_f, \theta_d) \right) \quad (6)$$

where n and n' are respectively the numbers of samples from the source and target domain. λ is the weight scalar of the loss. By implementing optimizers like SGD, Adam, or RMSProp, Equation 6 is expected to converge to a saddle point, and the labels of the target domain then can be predicted. It is noticeable that the DANN model can be used in either a semi-supervised or unsupervised way. The proposed method in this research is leveraging the unsupervised DANN with the Bi-LSTM feature extractor, which is discussed in the following section.

3. PROPOSED METHODOLOGY

3.1. Bearing degradation simulation dataset

As discussed in the previous section, the bearing DGT model could generate vibration signals based on real bearing geometry and operating conditions. However, it has to be noticed that the model does not explain the dynamic mechanism for the bearing failure and therefore cannot evolve with the bearing degradation. On the other hand, it is possible to have a

glance at this process using various Health Indicators (HI). Following the work of B. Wang et al. (2018), the bearing degradation process is estimated as an exponential function described as:

$$f(t) = \alpha e^{\beta t} + \gamma e^{\delta t} + \epsilon \quad (7)$$

where t is the time stamp, and $f(t)$ represents the HI extracted from the raw vibration signal. The other parameters from α to ϵ could be obtained using a nonlinear least squares method. The exponential curves should cover up a broad enough input feature space, containing most of the possible degradation trajectories, to provide abundant samples and instructive information for the follow-up learning process. The schematic of the simulation process is shown in Figure 4, where the real and simulated degradation processes are described with an increasing HI. t_p is the time stamp when the first prediction is made, and t_λ is the inspecting time. The solid red line indicates the real measurements and the dotted red line is the indicators that has not been measured.

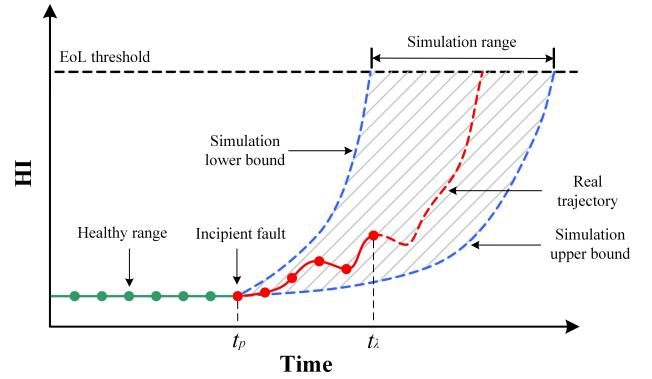


Figure 4. Simulation of degradation based on HI trajectories.

The baseline degradation trajectory is obtained based on bearing's L_{10} life. When a specific bearing is operated under certain speed and load, its L_{10} life can be calculated to infer the bearing reliability (Huang et al., 2007). The baseline exponential function could be fitted based on the End-of-Life (EoL) threshold and the L_{10} life value which is then extended for further simulation.

3.2. RUL prediction with Bi-LSTM based DANN

In this research, sensory vibration signals of bearings are collected as the input data. The flow chart of the proposed method is depicted in Figure 5, which is composed as follows. The source domain data includes the vibration signals from the simulation degradation datasets and the corresponding RUL scalar values as labels. The target domain data contains the vibration signals measured from the real bearing. Then these two types of information are passed to the proposed neural network.

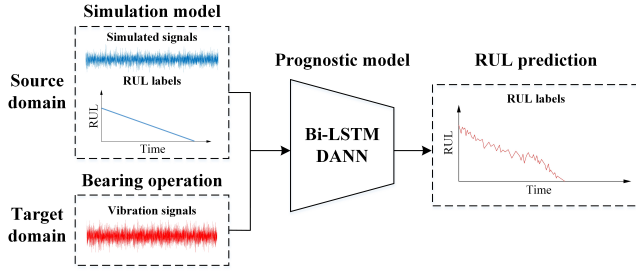


Figure 5. Flow chart of the proposed prognostic method.

Bi-directional Long-Short Term Memory (Bi-LSTM) layers are combined with Fully Connected (FC) layers to form the feature extractor in the DANN model. The FC layers are also used in the domain classifier and the label predictor with shared weights from the Bi-LSTM. A softmax and a linear activation function are applied respectively in the last layers of the domain classifier and the label predictor. In this implementation, the cross-entropy loss function is used on the domain classifier. Since the label predictor is dealing with the regression problem, the root mean square error loss function is selected for its optimization. The detailed architecture and the parameter selection of the proposed Bi-LSTM DANN can be found in Table 1.

Table 1. Architecture of the Bi-LSTM DANN.

| Feature extractor | | |
|-------------------|-----------------------|------------------------------|
| 1 | Input layer | Size: length of the signal |
| 2 | Bi-LSTM layer | Unit: 500 |
| 3 | Dropout layer | Rate: 0.2 |
| 4 | FC layer | Unit: 200, activation: Relu |
| 5 | Dropout layer | Rate: 0.2 |
| 6 | Feature output | Unit: 200, activation: Relu |
| Domain classifier | | |
| 7 | FC layer | Unit: 100, activation: Relu |
| 8 | Dropout layer | Rate: 0.2 |
| 9 | FC layer | Unit: 50, activation: Relu |
| 10 | Domain classification | Unit: 2, activation: Softmax |
| Label predictor | | |
| 11 | FC layer | Unit: 100, activation: Relu |
| 12 | Dropout layer | Rate: 0.2 |
| 13 | FC layer | Unit: 50, activation: Relu |
| 14 | Source regression | Unit: 1, activation: Linear |

Besides the layers mentioned above, the GRL is added between the last FC layer of the feature extractor and the first FC layer of the domain classifier, as discussed by Ganin et al. (2016). The weight of loss, λ , between the domain classifier and the label predictor is set to 1.0 in this work.

4. EXPERIMENTAL PART

4.1. IMS dataset

The proposed method is validated using the NASA IMS dataset including three subsets (Qiu et al., 2006). Each of them contains the vibration signals measured from run-to-failure bearing experiments as shown in Figure 6. Four (4) double row Rexnord ZA-2115 bearings were mounted on a shaft connected with an electric motor.

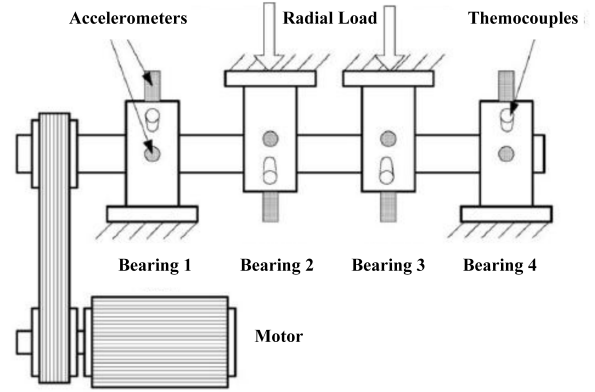


Figure 6. Test rig layout of IMS experiment.

Two accelerometers were installed on each bearing to acquire signals from x- and y-axes. With the sampling frequency of 20 kHz, each individual measurement lasted for 1 second and the recording interval was 10 min. In the released documents, each file represents one vibration signal. Four (4) bearings were observed as defected with different fault types and lifespans. The original End-of-Life (EoL) threshold of the experiment was defined based on the accumulated debris of a magnetic plug, which lacked of detailed measuring information. In order to quantify the bearing life from the vibration perspective, the EoL adopted in this research is reached when the acceleration overcomes 5g.

In this research, the prognostics is proceeded after the detection of an anomaly based on engineered features, which is assumed as the occurring of incipient fault. The faulty bearing information are listed in Table 2. The anomaly detection results are based on the research of C. Liu & Gryllias (2020).

Table 2. Faulty bearing information of IMS dataset.

| Bearing | Fault type | Measurement duration | Anomaly detection |
|------------------------|------------|----------------------|-------------------|
| Dataset 1 Bearing 3 | Inner race | 2156 files | Signal #1796 |
| Dataset 2 Bearing 1 | Outer race | 980 files | Signal #532 |
| Dataset 3 Bearing 3 | Outer race | 6323 files | Signal #5977 |

4.2. Comparative methods

IMS dataset has been used as a benchmark in various researches in the field of bearing prognostics. Four published prognostic methods based on this dataset are reviewed and used as the targets to compare against the proposed approach.

- *SVM* (Dong & Luo, 2013): The method firstly employs Principal Component Analysis (PCA) to extract features from the inputs and combines them with a SVM to conduct the prognostics.
- *LSTM* (Cheng et al., 2018): This method uses signal processing methods to extract classic engineered features from both time and frequency domain to feed a deep LSTM neural network.
- *PSW* (Qian et al., 2017): A hybrid model using Phase Space Warping (PSW) technology is used together with the Paris' law to estimate the RUL based on crack growth.
- *CNN* (R. Liu et al., 2019): A CNN model with a joint loss function is proposed which can simultaneously diagnose the bearing fault type and predict the RUL.

The results from these papers are gathered to compare with the proposed method. It should be noticed that, some of the published methods use supervised learning methods based on fleet-wise model training, which require prior knowledge about the degradation process. With the help of the bearing DGT, the prognostics is proceeded in an unsupervised way in this research, which needs only the bearing geometry and the operating conditions as the inputs.

Beside comparing with published results, it is also necessary to examine the effectiveness of the domain adaptation. Therefore the proposed model is also compared against the non-adapted method. Instead of the adversarial training during the modelling process, the non-adapted model is constructed based on the simulation and then is tested directly with the real data.

4.3. Evaluation metrics

In this research, the predicted RUL value, RUL_{pred} , can be calculated based on the normalized RUL in $[0, 1]$ and the actual inspecting time t_λ :

$$RUL_{pred} = \frac{RUL_{norm}}{1 - RUL_{norm}} t_\lambda \quad (8)$$

The performance of the prognostic model can be measured via several metrics (Saxena et al., 2010). Two error-based metrics, the Root Mean Squared Error (RSME) and the Mean Absolute Error (MAE) are commonly adopted in prognostic research to calculate the error between RUL_{act} and RUL_{pred} . Considering there are entirely N measurements, the errors can be described as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (RUL_{act}^i - RUL_{pred}^i)^2}{N}} \quad (9)$$

$$MAE = \frac{\sum_{i=1}^N |RUL_{act}^i - RUL_{pred}^i|}{N} \quad (10)$$

Error-based metrics can globally evaluate the predictions which equally weight the samples at different time stamps. In practice, a precise prediction at the time closer to the EoL is considered more valuable for decision making than at the starting period of the degradation. Therefore, the Cumulative Relative Accuracy (CRA) is used as an aggregate prediction accuracy metric, which is described as:

$$CRA = \sum_{i=1}^N \omega_i RA_i \quad (11)$$

where RA represents the relative accuracy at specific time stamp and ω is the corresponding weight. They can be calculated as follows:

$$RA_i = 1 - \frac{|RUL_{act}^i - RUL_{pred}^i|}{RUL_{act}} \quad (12)$$

$$\omega_i = \frac{i}{\sum_{i=1}^N i} \quad (13)$$

When the RA and CRA values are closer to 1, the model presents a good prognostic performance. Moreover, the α - λ accuracy is used to evaluate the RUL estimation at each individual time stamp. α represents the tolerance range of the true RUL, which is also denoted as the α -cone. t_λ indicates the inspecting time stamp, and λ is in the range $[0, 1]$ representing the fraction between the starting time stamp t_p and the EoL. The prediction at t_λ should be evaluated if it falls into the α -cone or not. t_λ can be expressed as follows:

$$t_\lambda = t_p + \lambda(EoL - t_p) \quad (14)$$

5. RESULTS

The proposed model is implemented with Python 3.6 and TensorFlow 1.9 on an Intel Xeon Gold 6140 (2.3 GHz) and NVIDIA Tesla P100 GPU. The bearing DGT models are constructed with an inner race and an outer race defect, respectively. Each model includes 100 degradation trajectories to cover all potential degradation modes and provide a wide enough latent feature space.

In order to fit the input shape of the Bi-LSTM layer, both the simulated and the real time sequences are reshaped as $[N, T, F]$ which are respectively the number of samples, time steps and features. During the training, the network is updated with the incoming target data sample N , which is set to 100. The experiments are repeated 5 times to reduce the influence of randomness, and the averaged metrics are recorded.

5.1. Simulation results

The simulated signals are generated based on the heretofore mentioned phenomenological model. It should be noticed that the amplitude parameter $q(t)$ in Equation 1 is selected based on the amplitude of the healthy signals before the anomaly happens. The real and the simulated signals of the IMS Dataset 2 Bearing 1 are shown in Figure 7.

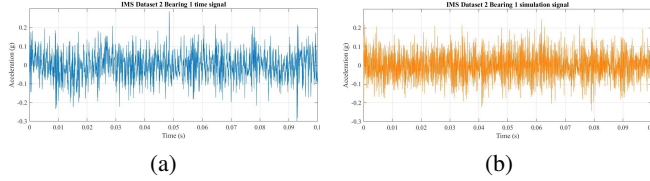


Figure 7. Real and simulated signals of IMS Dataset 2 Bearing 1 (a) real signal. (b) simulated signal.

5.2. Comparison against published results

As illustrated in the previous section, the comparison of RUL prediction results is firstly carried among the published state-of-the-art results. The prognostic results from the selected papers are listed in Table 3. It is noticeable that almost all the bearing prognostic papers with the IMS dataset only validate the models on the faulty bearings in the first two subsets. The metrics are calibrated using the number of measured samples based on Equation 8. In the table, the symbol “-” indicates that the results are not provided.

The results of the DGT-based DANN model are also presented for the two faulty bearing cases. With lower prediction errors and higher CRA values, the superiority of the proposed method can be attributed to the DGT-based simulation as well as the DANN’s domain generalization ability. In the prognostic tasks, classic deep neural networks extract features based on the limited number of measurements from a single degradation trajectory. These one-on-one methods could easily fall into the trap of a case-specified or overfitted model, which cannot be applied to other datasets. On the other hand, many existing deep learning methods treat prognostics as a supervised learning problem with training and testing datasets split from a single run-to-failure experiment. However, it is impractical to make such an assumption dealing with on-line dataflow under real prognostic scenarios.

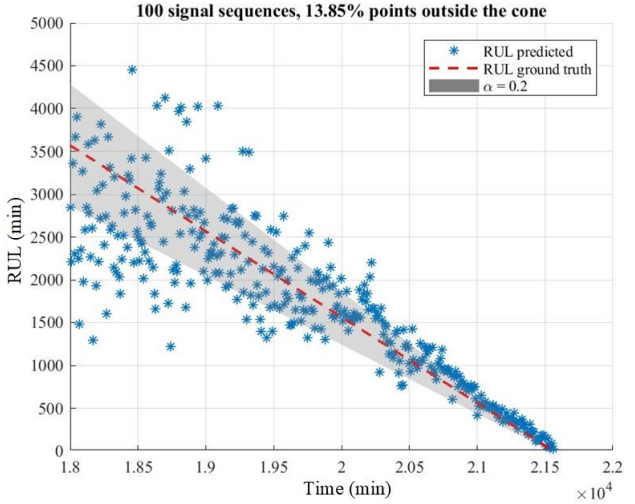


Figure 8. Prognostic results of Dataset 1 Bearing 3.

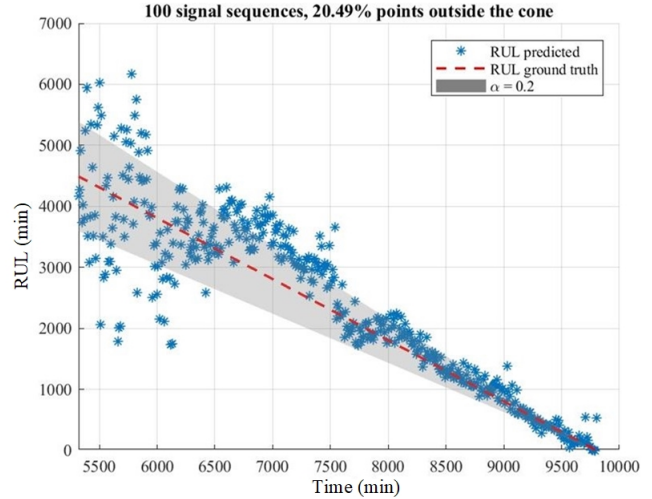


Figure 9. Prognostic results of Dataset 2 Bearing 1.

Based on the proposed method, the RUL predictions and the α - λ performance of Dataset 1 Bearing 3 and Dataset 2 Bearing 1 are shown in Figure 8 and Figure 9, respectively. In the first bearing case, it can be observed that the predictions before 1.9×10^4 minute ($\lambda = 0.27$) are decentralized from 1,200 to 4,500 minutes. The domain-invariant features

Table 3. Comparison of state-of-the-art prognostic models with the proposed method based on IMS dataset.

| Metric | Dataset 1 Bearing 3 | | | | | Dataset 2 Bearing 1 | | | | |
|------------|---------------------|------|-----|--------|--------|---------------------|--------|--------|--------|--------|
| | SVM | LSTM | CNN | PSW | DANN | SVM | LSTM | CNN | PSW | DANN |
| RMSE (min) | - | - | - | 547.20 | 511.80 | 429.60 | 364.20 | 615.55 | 857.40 | 349.20 |
| MSE (min) | - | - | - | 425.40 | 348.60 | - | 361.80 | - | 681.00 | 241.20 |
| CRA | - | - | - | - | 0.72 | 0.64 | - | - | - | 0.78 |

have not been fully captured from the real measurement due to the lack of training samples in the target domain. With the increase of available training data, the model’s performance is significantly boosted with the monotonic trend observed from the predictions, especially after 2.05×10^4 minutes ($\lambda = 0.68$). Generally, 13.85% of the predictions are outside the α -cone.

A similar performance is found in the second bearing case with a faster converge of the prediction starting at 7,500 minute ($\lambda = 0.49$). The proportion of the out α -cone predictions is 20.49%, as shown in Figure 9.

With the assist of simulation data, the physical information of bearing degradation is embedded in the source domain. The neural network training will further result in the emergence of domain-invariant and monotonic features, which can be used to track the bearing degradation. Since the training of the target domain is unsupervised, the proposed DGT-based DANN model is able to circumvent the labelling issue of the run-to-failure dataset. In practical applications, the available unlabelled measurements increase with the machine operating and could be used to enhance the target domain, thus leading to a more precise prediction.

5.3. Comparison against non-adapted models

In order to further validate the effectiveness of the domain adaptation, a comparative analysis has been performed between the proposed method and a non-adapted DGT model, which is only trained with the simulation data without the domain classifier. The simulation model construction keeps consistent with the proposed method using 100 generated signal sequences. The RUL prediction errors and the computational cost of Dataset 3 Bearing 3 are reported in Table 4. It should be noticed that the adversarial training process increases the training time compared to the non-adapted models. The corresponding prediction curves are depicted in Figure 10.

Table 4. Prognostic metrics of Dataset 1 Bearing 3.

| | DGT based-DANN | Non-adapted model |
|---------------------|----------------|-------------------|
| RMSE (min) | 445.59 | 790.22 |
| MAE (min) | 328.24 | 715.30 |
| CRA | 0.68 | -2.09 |
| Training time (min) | 16.78 | 10.32 |

Compared to the DANN method, the non-adapted model fails to adapt its outputs to the real data, which gives rise to high prediction errors. Although the simulation dataset could provide enough training samples to the neural network, the shift between the simulation and real is still considerable since the DGT cannot perfectly replicate the degradation process,

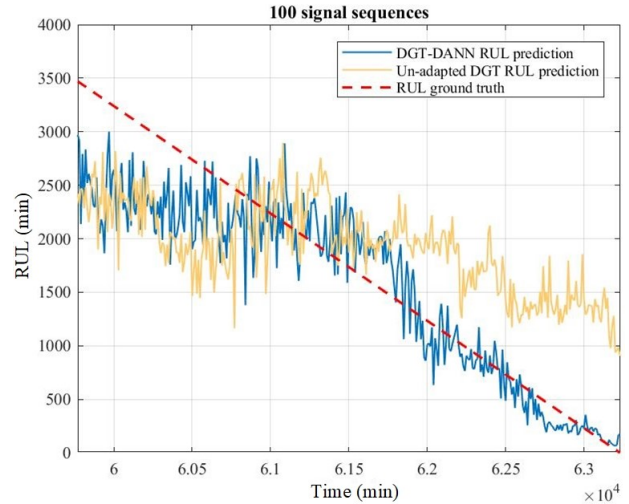


Figure 10. Prognostic results of Dataset 3 Bearing 3.

which leads to the higher prediction errors and less monotonic RUL curve.

On the other hand, the DGT-based DANN method shows the promising results of transferring physical information to the deep learning framework via domain adaptation. The results reveal that the proposed DGT-based DANN model is a sound approach to be used in bearing prognostics.

6. CONCLUSION

Combining deep learning with a digital twin is one of the key opportunities in the future-oriented smart industry. In this paper, a bearing DGT model is married with a deep neural network in the domain adaptation frame to fulfill a prognostic task. At the heart of this prognostic approach lie two critical techniques: the bearing vibration model and the domain adversarial neural network. The experimental results based on the IMS dataset validate the efficacy of the DGT-based DANN method. Among the state-of-the-art prognostic results, the proposed method exhibits dominant performance with lower prediction errors and high relative accuracy. The comparison between the adapted and non-adapted models confirms the virtues of the DANN. Besides these encouraging results, the method provides a new route to handle the unsupervised learning problem in practice, which has plagued the prognostic community for a long time. The proposed method not only contributes as a data driven framework for the bearing prognostic task, but also explores a new path to synchronize the physics-based simulation and the deep learning. More sophisticated simulation models could be used as alternatives for the training data augmentation, which might lead to a better prediction. From the industry point of view, the deployment of the proposed model will significantly ease the reliance on real operational data with lower cost.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of the China Scholarship Council, Flanders Make the strategic research centre for the manufacturing industry in the context of DGTwin Prediction project, Flemish Government under the “Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme, and the Flemish Supercomputer Center (VSC) for the computation resource.

REFERENCES

- Antoni, J. (2007). Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions. *Journal of Sound and vibration*, 304(3-5), 497–529.
- Antoni, J., & Randall, R. (2003). A stochastic model for simulation and diagnostics of rolling element bearings with localized faults. *J. Vib. Acoust.*, 125(3), 282–289.
- Chen, Z., He, G., Li, J., Liao, Y., Gryllias, K., & Li, W. (2020). Domain adversarial transfer network for cross-domain fault diagnosis of rotary machinery. *IEEE Transactions on Instrumentation and Measurement*.
- Cheng, Y., Zhu, H., Wu, J., & Shao, X. (2018). Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks. *IEEE Transactions on Industrial Informatics*, 15(2), 987–997.
- da Costa, P. R. d. O., Akcay, A., Zhang, Y., & Kaymak, U. (2020). Remaining useful lifetime prediction via deep domain adaptation. *Reliability Engineering & System Safety*, 195, 106682.
- Dong, S., & Luo, T. (2013). Bearing degradation process prediction based on the pca and optimized ls-svm model. *Measurement*, 46(9), 3143–3152.
- Ganin, Y., & Lempitsky, V. (2014). Unsupervised domain adaptation by backpropagation. *arXiv preprint arXiv:1409.7495*.
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... Lempitsky, V. (2016). Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 17(1), 2096–2030.
- Gao, Y., Liu, X., & Xiang, J. (2020). Fem simulation-based generative adversarial networks to detect bearing faults. *IEEE Transactions on Industrial Informatics*, 16(7), 4961–4971.
- Gryllias, K. C., & Antoniadis, I. A. (2012). A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments. *Engineering Applications of Artificial Intelligence*, 25(2), 326–344.
- Ho, D., & Randall, R. (2000). Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical systems and signal processing*, 14(5), 763–788.
- Huang, R., Xi, L., Li, X., Liu, C. R., Qiu, H., & Lee, J. (2007). Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. *Mechanical systems and signal processing*, 21(1), 193–207.
- Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 107, 241–265.
- Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to rul prediction. *Mechanical Systems and Signal Processing*, 104, 799–834.
- Liu, C., & Gryllias, K. (2020). A semi-supervised support vector data description-based fault detection method for rolling element bearings based on cyclic spectral analysis. *Mechanical Systems and Signal Processing*, 140, 106682.
- Liu, R., Yang, B., & Hauptmann, A. G. (2019). Simultaneous bearing fault recognition and remaining useful life prediction using joint-loss convolutional neural network. *IEEE Transactions on Industrial Informatics*, 16(1), 87–96.
- McFadden, P., & Smith, J. (1984). Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of sound and vibration*, 96(1), 69–82.
- Qian, Y., Yan, R., & Gao, R. X. (2017). A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mechanical Systems and Signal Processing*, 83, 549–567.
- Qiu, H., Lee, J., Lin, J., & Yu, G. (2006). Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of sound and vibration*, 289(4-5), 1066–1090.
- Randall, R. B., Antoni, J., & Chobsaard, S. (2001). The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical systems and signal processing*, 15(5), 945–962.
- Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for offline evaluation of prognostic performance. *International Journal of Prognostics and health management*, 1(1), 4–23.
- Sobie, C., Freitas, C., & Nicolai, M. (2018). Simulation-driven machine learning: Bearing fault classification. *Mechanical Systems and Signal Processing*, 99, 403–419.
- Wang, B., Lei, Y., Li, N., & Li, N. (2018). A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Transactions on Reliability*.
- Wang, Q., Michau, G., & Fink, O. (2019). Domain adaptive transfer learning for fault diagnosis. In *2019 prognostics and system health management conference (phm-paris)* (pp. 279–285).

BIOGRAPHIES

Chenyu Liu received the B.S. and M.Sc degree in aircraft manufacturing engineering and solid mechanics both from Northwestern Polytechnical University, China. He joined the Noise and Vibration Research Group in the Department of Mechanical Engineering, KU Leuven, Belgium as a PhD researcher in 2017. His research interests include data-driven based condition monitoring, rotating machinery prognostics, and machine learning and deep learning application.

Alexandre Mauricio received his MSc. degree in Mechanical Engineering from the Faculty of Engineering of the University of Porto (FEUP) in 2017. He is currently a PhD student at the Department of Mechanical Engineering of KU Leuven, Belgium. His main research interests focus on the diagnostics of rotating machinery operating under time varying conditions.

Junyu Qi received his B.S. and M.S. degree in Mechatronics from Shanghai Maritime University, Shanghai, China, in 2012 and from Friedrich-Alexander-University Erlangen-Nuremberg, Germany, in 2016, respectively. He is currently a PhD student in the Department of Mechanical Engineering of KU Leuven, Belgium. His research interests focus on condition monitoring of rotating machinery and particularly in prognostics.

Dandan Peng was born in Shaanxi, China, in 1992. She received the B.S. and M.S. degrees in mechanical engineering from the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China, in 2016 and 2019, respectively. She is currently working toward the Ph.D. degree in mechanical engineering from Catholic University of Leuven, Leuven, Belgium. Her research interests include Hilbert–Huang transform, convolutional neural network, machinery condition monitoring, and fault diagnosis.

Konstantinos Gryllias holds a 5 years engineering diploma degree and a PhD degree in Mechanical Engineering from National Technical University of Athens, Greece. Since October 2014 he holds an assistant professor position on vibro-acoustics of machines and transportation systems at the Department of Mechanical Engineering of KU Leuven, Belgium. He is also the manager of the University Core Lab Dynamics in Mechanical & Mechatronic Systems DMMS-M of Flanders Make, Belgium. His research interests lie in the fields of condition monitoring, signal processing, prognostics and health management of mech. & mechatronic systems.