Long short-term memory based semi-supervised encoder–decoder for early prediction of failures in self-lubricating bearings

Vigneashwara PANDIYAN¹,*, Mehdi AKEDDAR¹, Josef PROST², Georg VORLAUFER², Markus VARGA², Kilian WASMER³

¹ Laboratory for Advanced Materials Processing (LAMP), Swiss Federal Laboratories for Materials Science and Technology (Empa), Thun CH-3602, Switzerland
² AC2T research GmbH, Wiener Neustadt 2700, Austria

Received: 27 September 2021 / Revised: 09 November 2021 / Accepted: 04 December 2021
© The author(s) 2021.

Abstract: The existing knowledge regarding the interfacial forces, lubrication, and wear of bearings in real-world operation has significantly improved their designs over time, allowing for prolonged service life. As a result, self-lubricating bearings have become a viable alternative to traditional bearing designs in industrial machines. However, wear mechanisms are still inevitable and occur progressively in self-lubricating bearings, as characterized by the loss of the lubrication film and seizure. Therefore, monitoring the stages of the wear states in these components will help to impart the necessary countermeasures to reduce the machine maintenance downtime. This article proposes a methodology for using a long short-term memory (LSTM)-based encoder–decoder architecture on interfacial force signatures to detect abnormal regimes, aiming to provide early predictions of failure in self-lubricating sliding contacts even before they occur. Reciprocating sliding experiments were performed using a self-lubricating bronze bushing and steel shaft journal in a custom-built transversally oscillating tribometer setup. The force signatures corresponding to each cycle of the reciprocating sliding motion in the normal regime were used as inputs to train the encoder–decoder architecture, so as to reconstruct any new signal of the normal regime with the minimum error. With this semi-supervised training exercise, the force signatures corresponding to the abnormal regime could be differentiated from the normal regime, as their reconstruction errors would be very high. During the validation procedure for the proposed LSTM-based encoder–decoder model, the model predicted the force signals corresponding to the normal and abnormal regimes with an accuracy of 97%. In addition, a visualization of the reconstruction error across the entire force signature showed noticeable patterns in the reconstruction error when temporally decoded before the actual critical failure point, making it possible to be used for early predictions of failure.

Keywords: predictive maintenance; in-situ sensing; long short-term memory (LSTM); encoder–decoder; wear monitoring; tribology

1 Introduction

In most real-world applications, wear occurs owing to partial solid–solid contact between sliding surfaces over a prolonged period. Such wear may be responsible for the catastrophic failures of machines made of moving parts. Hence, full visibility of the machine conditions enables machine manufacturers to adopt effective maintenance plans, such as preventive, predictive, or reactive plans, so as to reduce downtime and increase productivity [1, 2]. Predictive maintenance techniques improve equipment maintenance efficiency...
relative to other maintenance plans by performing cost-effective maintenance only when necessary [3, 4]. In other words, the current state of the equipment or its components is assessed to predict when the maintenance should be performed, leading to a longer equipment uptime [5]. The principle of predictive maintenance is to continuously monitor data correlated to the machinery’s state through analytics to recognize patterns revealing that failure is likely to occur [6–9]. The mitigation of these failures (those forecasted to occur) can then prevent the machine from unplanned downtime.

Self-lubricating bearings have an increased lifetime and improved wear properties relative to those of other bearing types. A constant, built-in lubrication film is generated in self-lubricating bearings; this film significantly reduces friction and direct solid contact [10]. Apart from the increased lifetime, these bearings also have the benefit of minimum maintenance as they do not require any greasing, making them an economical choice for reduced environmental impacts [11, 12]. Self-lubricating bearings based on polymer composites are fabricated with uniformly distributed holes or grooves on the base material surfaces. The grooves or holes are inserted with a solid polymer compound, generally a graphite base [13–15]. The surface area of the lubricant covers approximately 25% to 30% of the total bearing surface. The geomantic arrangements of the holes or grooves are arranged so as to cover the entire sliding surface in the movement direction [16, 17]. During the initial running-in operation of the two mating surfaces, the lubricant in the holes/grooves is released between the interfaces, thereby filling up the asperities of the two surfaces. Consequently, the solid lubricant establishes a continuous lubricating film between the contact sliding surfaces, and prevents metallic contact between the moving parts. The lubricating film is constantly replenished by lubricant particles emerging from the lubricant reservoir [14]. The solid lubricant also acts as a reservoir by absorbing foreign particles and dirt into the lubricant holes/grooves/reservoir. These embedded particles displace a similar amount of lubricant, which is then available for lubrication of the bearing surfaces. This absorption of unwanted particles and release of solid lubricant make the operation reliable and maintenance-free. Nevertheless, bearing failures can occur after prolonged use and remain one of the major causes of machinery breakdowns, making the detection and diagnosis of mechanical faults during their operation crucial [11, 18].

The best possible methodology for ensuring the quality of sliding surfaces is a direct visualization of the wear marks on them. However, such visualization techniques come with a high cost as, in most cases, the inspection cannot be done when the component is operational, owing to limited accessibility to the critical regions of the sliding surfaces [19, 20]. Under such circumstances, the process must be stopped before the component is removed and taken offline for inspection, resulting in machine downtime. Conducting such a laborious time-discrete offline inspection will hamper the overall productivity. A recent trend in the digital age for estimating the wear state is to interpret and identify patterns in the signatures of sensors with an indirect correlation with the surface state in the contact zone using machine learning (ML) algorithms [21–23]. Making decisions using ML algorithms has two advantages: First, the component’s wear is continuously monitored, unlike discrete offline inspection; second, the workforce and downtime knowledge are available for forecasting productivity and planning maintenance actions.

Orhan et al. [24] demonstrated that the progressive failures occurring on rolling element bearings can be determined early based on a spectral analysis of vibration sensor data. Prieto et al. [25] developed a diagnosis methodology by using a hierarchical neural network structure to classify six different bearing operating regimes according to vibration sensor data. Sadegh et al. [26] reported that the different lubrication conditions in journal bearings could be successfully determined using artificial neural networks and genetic algorithms. In terms of sensors, the acoustic emission signals captured from the contact zones of bearings have been combined with ML methods to detect anomalies [26–29]. Moder et al. [30] used neural networks and logistic regression to successfully classify lubrication regimes in a hydrodynamic bearing based on torque sensor signals. In contrast, Prost et al. [31] trained a random forest classifier on lateral force data corresponding to each cycle of an oscillating
Friction self-lubricating bearing set up to distinguish between four states of operation. In addition, the identification of degradation in bearings using ML models has been extensively researched [32–34]. Similar to traditional ML algorithms that work on sparse statistical features, convolutional neural networks have also been established for monitoring rolling bearings based on raw sensor data [35–37]. Narendiranath Babu et al. [38] reported using deep neural networks to classify the vibration data of the normal condition and five abnormal conditions in journal bearings. Sensing techniques such as thermal imaging [39, 40] have been combined with ML to classify the states of equipment. Based on the literature survey, we can conclude that if sensors can carry the representative signal information of the bearing state, any state-of-the-art ML algorithm can decode the patterns for monitoring the bearing conditions [19]. However, most ML algorithms treat the sensor data as a vector for monitoring purposes, and algorithms treating the sensor data as a sequence are very sparse [41]; this gap will be bridged in this work.

Unlike traditional feed-forward networks, recurrent neural networks (RNNs) do not take a fixed input, and do not process all vector data simultaneously [42]. In fact, an RNN reuses a computational unit that takes data sequentially and at each step of its calculations [43, 44]. When processing a sequence, the output at each step is combined with the next element from the sequence, and used as the next input. Therefore, another output is produced until the sequence is completed. As a result of this sequential processing, the RNN learns the temporal relationship between the elements in the sequence, which can be further exploited for other suitable downstream tasks [45]. Long short-term memory (LSTM) networks represent a type of RNN, and are capable of learning and holding more long-term temporal contextual information in a sequence compared to “vanilla” RNNs [46, 47]. Apart from holding the information for a long time, they are not prone to vanishing and exploding gradients, as in vanilla RNNs [48].

The key elements inside an LSTM block are the cell state and hidden state [49, 50], as shown in Fig. 1. The cell state is responsible for sharing and transferring information along the entire sequence chain [50].

However, while processing each element in the sequence, the cell states are altered by gates containing sigmoid and tanh activations inside the LSTM cell [51, 52]. The tanh function helps in regulating the network by squishing values between −1 and 1, and the sigmoid function helps to control the information to be kept (1) and forgotten (0) [51]. The hidden state is basically the output of the current element in the sequence after the gating operation, and is combined with the next element for computation [51, 52].

A major concern in a failure-related system is that a large amount of data represents a good process, and much less data are available presenting a flawed process or failure. Hence, the dataset between normal and abnormal processes is very unbalanced, resulting in increased biasing of the model during training. Furthermore, with the inaccessibility to the process zone, the correct labelling of the wear state across the full dataset is cumbersome. In this work, owing to the imbalanced dataset, we transform a classification problem with multiple classes into a binary classification problem using semi-supervised ML learning, where the model is trained with data whose nature is familiar to the machine operator knowledge-wise, i.e., a normal regime. The principle of a semi-supervised ML algorithm is to familiarize the model with sensor signatures corresponding to the desired class. In addition, the normal regime can be differentiated without retraining when data from a new class are included. Pandiyan et al. [53] developed a monitoring strategy for microphone data by using semi-supervised learning and a variational auto-encoder network. Compared to the latter, this current work differs in two ways, and improves the previously proposed methodology. First, it uses a force signature corresponding to each oscillation cycle to train the
encoder–decoder model based on the LSTM. Second, based on the temporal reconstruction loss of the model across each localized cycle in the whole dataset, the onset of a failure mechanism can be identified before the failure occurs, providing a novel result.

The remainder of this paper is organized into five sections. Section 1 summarizes the self-lubricating bearing design, monitoring technologies, and LSTM method. Section 2 presents the oscillating tribometer setup, data collection, and pre-processing data strategy. Section 3 provides information regarding the LSTM-based encoder–decoder architecture and training methodology used on the force signature acquired during the experiment. Section 4 presents the detection results from the failure regimes using the trained LSTM model. Finally, the findings of our contributions and future research directions are presented in Section 5.

2 Experimental setup

The experiments were conducted on a transversely oscillating tribometer setup described in more detail in previous publications [31, 53]. The tribological pairing consisted of a shaft made of hardened and polished Cr steel (100Cr6, \( \Omega \) 24 mm) sliding in tight contact inside a self-lubricating bronze bushing, as shown in Fig. 2.

The bushing was equipped with macro deposits of polymer lubricant, from which the lubricant was distributed over the sliding contact. The tribometer setup was equipped with a variety of sensors. The present study focused on interpreting and correlating force sensor data with failure regimes. An axial force sensor (HBM U9C, Germany) was mounted parallel to the moving axis of the shaft, and the normal force was recorded by a load cell with a higher load capacity (HBM U2B, Germany). The sensor data were acquired using in-house developed data acquisition and control software at a rate of 5 kHz. The detailed specifications of the force sensors are listed in Table 1. The oscillating movement of the shaft was recorded using a linear inductive position sensor (Turck Li300P0-Q17LM0-LiU5X2, Germany). The bearing temperature was measured using a type-K thermocouple with a diameter of 0.5 mm mounted inside a drilled hole on the bearing’s front face, as shown in Fig. 3. The force data were acquired from experiments performed by inducing the shaft movement at a nominal frequency of 1 Hz and stroke length of 50 mm. The experiment was designed to run with a specific bearing load of 8 N/mm², corresponding to a radial (normal) load of 6 kN. To establish a smooth run-in, the load was gradually increased during the first 1.5 h of the experiment.

Fig. 2 Experimental setup to simulate the wear in the self-lubricated bearing.
The experiments were run uninterrupted until one of the two stop criteria was reached, representing threshold values on the measured lateral force and sample temperature. The thresholds were set to $\pm 3.5 \text{kN}$ for the uncorrected lateral force signal and $150 \, ^\circ\text{C}$ for the bearing temperature. Datasets from four experiments were available for this study. The typical running times for all of the experiments in the study were between 10 and 16 h or approximately 36,000–57,000 cycles.

### 3 Proposed ML methodology

The primary objective is to detect an abnormal regime (running-in and anomaly are addressed together as an abnormal regime throughout this work) and the onset of failure, based on force signatures in the self-lubricating bushings when operational. The running-in and anomalies (abnormal regime) can be isolated from the normal operating regimes using an encoder–decoder architecture, where the network familiarizes reconstructing only the signatures of the normal regime, as they are only trained on them. The trained encoder–decoder model then will fail to achieve a good reconstruction when encountering signals corresponding to the abnormal regime. From the magnitude of the reconstruction loss, we can detect that the signals are not part of the normal regime.

A schematic of the proposed methodology is presented in Fig. 4.

The training strategy for the encoder–decoder architecture consists of three parts. The first part is dataset preparation. For reduced computations, the force signatures of each cycle are downsampled to 100 data points for the entire experimental data, such that the signature envelopes the shape of the cycle. The sequential vector lengths of 100 data points are the input for the LSTM-based encoder–decoder model for training and testing in real-time deployment. The accuracy of the trained encoder–decoder model primarily depends on the choice of the dataset distribution it is trained with. As the training dataset should be statistically different from the abnormal regime signature, the force signatures corresponding to each cycle representing the normal regime are separated from the entire dataset. The remaining

| Sensor          | Model   | Number of channels | Sampling rate | Sensor working range               | Acquisition mode |
|-----------------|---------|--------------------|---------------|------------------------------------|------------------|
| Axial force     | HBM U9C | 1                  | 5 kHz         | up to 10 kN 24 bit ADC             | Continuous       |
| Normal force    | HBM U2B | 1                  | 5 kHz         | up to 200 kN 24 bit ADC            | Continuous       |

Fig. 3  Transversally-oscillating tribometer setup integrated with the accelerometer, force, thermocouple, and acoustic emission sensor.
dataset represents the abnormal regime. The dataset representing the normal regime is split stochastically into three portions containing 60%, 20%, and 20% of the data, respectively. The second part of the training strategy includes training the actual LSTM model with supporting libraries, optimizers, training parameters, and finally, computing the threshold. The 60% portion of the dataset representing the normal regime during the previous step is used to train the model, which is verified based on the visualization of the loss curves. At the end of the training, the second portion (20%) of the normal regime's dataset is passed to the model, and the reconstruction loss distribution is computed. Next, a threshold value is computed from the reconstruction loss distribution (the sum of the mean and three standard deviations). The third and final part involves testing the model and confirming its performance. The third portions (20%) of the normal and whole abnormal datasets (including both running-in and anomaly) are passed into the trained model, and the reconstruction loss is computed. The insight behind such an analysis is that for the signals in normal regimes, the LSTM-based encoder–decoder model will produce a reconstruction loss lower than the threshold, as it will have learned the patterns. However, owing to their unfamiliarity, abnormal regime signals would produce a higher reconstruction loss than the threshold, from which they can be easily identified as outliers.

3.1 Dataset

This work focused on understanding the distribution of the force data from the oscillating trials of the normal regime to identify the abnormal regime. The axial force signals were acquired at a rate of 5 kHz, and time-series data corresponding to each cycle were stored for offline processing. The oscillating nature of the setup created periodic patterns across each cycle, as shown in Fig. 5. The zero position of each cycle was triggered using the zero-crossings of a positional encoder trigger embedded in the tribometer setup. The axial force signal representing each cycle was downsampled to 100 data points using linear interpolation such that these 100 data points represented each cycle’s envelope, as shown in Fig. 5.

Figure 6(a) shows the overall evolution of the downsampled force signal during one of the experiments, where the red curve represents the running-in, the green curve represents the normal regime, and the blue curve represents the anomaly. In Fig. 6(a), the differences between the running-in and the degradation of the bearing from the normal state of operation can be seen as distortions of the individual cycle shapes. Figure 6(b) compares the
downsampled force signals between the operational regimes for running-in, normal, and anomaly across ten randomly selected cycles.

In total, datasets from four experiments were available for training and testing the LSTM-based encoder–decoder architecture, as shown in Table 2. Based on the number of cycles \(m\), a matrix of size \(m \times 100\) was computed for each experimental trial,

| Experiment | Total number of cycles before failure | Variance in the dataset |
|------------|--------------------------------------|-------------------------|
| Experiment 1 | 44,265 | 0.83 |
| Experiment 2 | 38,605 | 0.89 |
| Experiment 3 | 57,516 | 0.86 |
| Experiment 4 | 35,388 | 0.80 |
and the trial was performed using the parameters discussed in Section 2.1. The matrix dataset for each experiment was heuristically segregated into running-in, normal, and anomaly data. Furthermore, the running-in and anomaly cycles were grouped as abnormal regimes. Thus, a dataset representing normal and abnormal regimes was prepared.

### 3.2 Visualization of wear

After reaching the end criteria (either by exceeding ±3.5 kN for the uncorrected lateral force signal or 150 °C for the bearing temperature), the bearings were examined by imaging the sliding surface with a microscope, as shown in Fig. 7. The microscope used was an Olympus SZX16 with a continuously variable focus, as equipped with a JENOPTIK ProgRes C5 microscope camera. With a closer examination of the functional surface of the bearing, we were able to observe the different wear mechanisms that occurred during the operational time, including scuffing marks, scratch marks, formation of wear debris at the interface edges between the lubricant reservoir and bushing body, and evidence of debris particles becoming embedded temporarily inside the solid lubricant reservoir. Most of the mechanisms were observed to occur simultaneously, with varying severities. The scratch marks could have resulted from short-time metal-to-metal contact when the lubrication film failed to replenish and/or third-body abrasion when the wear debris was dislodged between the moving surfaces.

### 3.3 LSTM architecture and training

LSTMs have been used to model temporal ordering relationships between data sequences such as text, audio, and time-series data [54–56]. In this work, the characteristics of the LSTM for remembering information in sequences were combined with the encoder–decoder architecture to compress this sequential representation and reconstruct it again, aiming to identify abnormal regimes. An illustration of the LSTM-based encoder–decoder architecture used in this study is shown in Fig. 8.

A four-layer stacked LSTM architecture was selected to build the encoder–decoder model, with two layers corresponding to the encoder (E) and the remaining two layers for the decoder (D). The encoder consisted of two LSTM blocks that read the input sequence of length 1 × 100 × 1 in a step-by-step manner. When processing the entire sequence, the representation learned by the recurrent units of the encoder returned a fixed-length vector of size 1 × 128. The fixed-length vector, otherwise called the latent space, was then passed as an input to the recurrent units based on the LSTM, which reconstructed the original input. This part of the architecture also had two LSTM blocks forming the decoder. Thus, the overall configuration of the encoder and decoder

![Fig. 7 Macro images of the worn bushings after the experiments.](image-url)
accomplished reading, encoding, representing, decoding, and finally reconstructing any input sequence. The model’s performance was computed based on its ability to reconstruct the temporal sequence distribution it was trained with; in our case, this was the normal regime force data. The encoder–decoder model was developed using built-in functions from the PyTorch library [57], and was trained on a hardware-accelerated graphical processing unit environment, namely the NVIDIA® Titan, using the training parameters in Table 3. During training, the signature of the normal operational regimes of each cycle was passed into the network and allowed to reconstruct. The difference between the original and reconstructed signals was computed as a loss, i.e., the mean square error. This loss was backpropagated to adjust the weights of the recurrent units for better reconstruction in subsequent iterations. The LSTM autoencoder model had 990,465 trainable parameters, was trained using the Adam optimizer [58] with a learning rate of 0.001, and was stopped after 100 epochs.

### Table 3 Parameters used for training the encoder–decoder network.

| Training condition/parameter | LSTM encoder–decoder |
|-----------------------------|----------------------|
| Optimizer used for training | “SGD”                |
| Learning rate               | 0.001                |
| Recurrent architecture      | Encoder–decoder architecture using 4 layers of LSTM |
| Epochs                      | 150                  |
| Size of the batch           | 1                    |
| Shuffle                     | Every-epoch          |
| Training data set (Normal regime) | Force signal of each cycle corresponding to normal operational regime downsampled to 100 continuous datapoints |
| Network input Loss          | 1 × 100 × batch size |
| Computational hardware      | GeForce Titan        |
| Dropout rate                | 0.1                  |
| Deep learning library       | PyTorch               |
| Trainable parameters        | 990,465              |

**Fig. 8** Illustration of the operation of the LSTM-based encoder–decoder network with \( X, H \) and \( F_C \) representing input state, hidden state, and fully connected layer.

**4 Failure prediction**

**4.1 Results and discussion**

Figure 9 shows the training and validation loss plots after training the LSTM-based encoder–decoder model for 150 epochs using the force signal envelopes representing cycles corresponding to the normal regime. In Fig. 9, based on the saturation of the loss curves, it is evident that the model learns the temporal distributions of the envelopes corresponding to the dataset (60% of the normal regime signals) on which it is trained, for better reconstruction. There is a significant decrease in the loss until 50 epochs, after which it becomes saturated with little improvement, indicating that the model has learned the distribution space. The validation loss plot also shows a similar trend to the training curve, indicating that the model prediction is robust.
From the trained model, a reconstruction threshold must be computed so that the model can flag the incoming signal as a normal or abnormal regime based on its ability to reconstruct, i.e., by comparing it with the threshold. The threshold is computed on 20% of the normal regime dataset that the model was unfamiliar with during the training. The reconstruction loss distributions for the normal regime envelopes from the four experiments are shown in Fig. 10. The distribution of the reconstruction loss lies between 0 and 2,000. A threshold value of 553.09 is calculated from the distribution on the upper value, based on Eq. (1).

Threshold = mean (μ) + 3 × standard deviation (σ) \hspace{1cm}(1)

In other words, if the model has poor reconstruction for an input signal, i.e., the reconstruction loss is greater than the threshold value of 553.09, it is assumed that the signal does not belong to the distribution of the normal regime dataset, and it is considered as an abnormal regime. Figures 11(b) and 11(c) show the reconstruction loss distribution for each cycle of the force signal in the abnormal regime datasets, namely the running-in and anomaly regimes, respectively. These distribution plots confirm that the distribution of the abnormal regime is higher than the threshold value of 1,346.4. However, as shown in Fig. 11(b), most reconstruction losses are well below the threshold. Consequently, we can postulate that the proposed LSTM-based autoencoder architecture can differentiate normal regimes from other undesirable regimes. We verified our approach on a test dataset consisting of 0.1 million cycles from the four experiments covering normal and abnormal regimes to obtain statistical confidence. The trained LSTM autoencoder model correctly classified 97% of the signals from the normal and abnormal regimes when compared with the ground-truth labels.

Figure 12 shows the trained model’s reconstruction of the force signals corresponding to the normal, anomaly, and running-in operational regimes. In the case of the normal regime, as depicted in Fig. 12(b), the reconstruction of the signal envelopes are excellent as the distribution is learned during training by the recurrent model. However, the reconstruction is poor in the running-in and anomaly regimes owing to the unfamiliarity, as shown in Figs. 12(a) and 12(c), respectively.

4.2 Early prediction of failure

The failure of a bearing owing to wear mechanisms does not generally occur instantaneously, but occurs over time. Thus, the wear progresses until the parts cannot sustain the applied load, leading to failure. Therefore, there should be a transition of the wear mechanism from the normal regime to the anomaly regime where the failure occurs. By carefully selecting the normal regime for training with minimal overlap with the anomaly regime, the normal regime can be segregated, as discussed in Section 4.1. The trained LSTM model can still reconstruct the signal in the transition zone owing to the overlap in the latent space representation between the normal regime and failure point. However, the reconstruction will be poor and generally higher than the threshold, i.e., the uncertainty of the model will increase. As the distribution in the transition zone moves more
toward failure, there is a significant decrease in the reconstruction ability of the model. Nevertheless, the temporal evolution of the reconstruction loss provides a view regarding the onset of the progressive wear mechanism. The same can be applied to the transition from the running-in regime to the normal regime. However, the reconstruction loss is minimized in this case, as the transition zone moves more toward a normal regime. The complete sequence of the downsampled force signals from Experiment 2 (Table 2) is presented in Fig. 13(c). Figure 13(b) shows the temporal evolution of the reconstruction loss across each cycle, from running-in to the failure regime. As cycles 8,000 to 18,000 were used to train the LSTM model, they were not detected as an anomaly, as inferred from the reconstruction loss patterns in Fig. 13(b). The temporal visualization of the reconstruction error in the transition zone is interesting, although it is well above the reconstruction threshold. We can see a positive drift in the magnitude of the reconstruction loss, as shown in Fig. 13(a). The slope of this drift, or the number of times there is a positive increment across the cycles, can be used to forecast the failure point. For Experiment 2 and all other experiments, evidence on the reconstruction loss suggests that wear mechanisms start at approximately 3,000–4,000 cycles prior to the actual failure point, giving a time window for maintenance action during the application.

5 Conclusions

This work addressed two tasks: First, segregating an abnormal state from a normal state; and second,
providing a methodology for early failure detection in self-lubricating bearings, even before it occurs. A four-layer stacked LSTM unit with a built-in encoder–decoder architecture was trained with a semi-supervised “model” to detect the different regimes to allow for predictive maintenance. The experiments were performed using a transverse oscillating tribometer setup. The training of the LSTM model with only the normal regime signal ensured that it only learned this discrete distribution. During the validation procedure of the proposed LSTM-based encoder–decoder model, the model predicted force signals corresponding to the normal regime and failure regime with an accuracy of 97%. The overall conclusions of this work can be summarized as follows.

1) With a relatively simple downsampled envelope representation (100 datapoints/cycle), the force signatures can successfully identify the operational regimes.

2) The abilities of the trained LSTM-based encoder–decoder model suggest that normal and abnormal operational regimes can be classified even when the dataset is very unbalanced. The semi-supervised training also opens up the possibility of using ML to classify regimes with minimum knowledge, and/or where data collection is expensive and cumbersome.

3) The postmortem visualization of the inner surface of the bearing reveals that multiple wear mechanisms can occur, and irrespective of the wear mechanism, the semi-supervised ML framework can separate normal operational regimes from failures, i.e., abnormal regimes.

4) The visualization of the reconstruction error across the cycles of each experimental trial offline shows

---

**Fig. 13** Early prediction of failures by analysis of the trend in the reconstruction loss on Experiment 2 (for better understanding, it is advised to consider the figure from bottom to top).
notable drift patterns temporally before the actual critical failure point. These can be used for early predictions or forecasting of the failures in future online analyses. The onset of failure regimes starts around 3,000–4,000 cycles before the actual failure for the experimental parameters used.

The performance of the LSTM framework suggests that the methodology can be utilized for predictive maintenance for any cyclic tribological experiment or machine component. This work can also be extended to other types of bearings or machine components. The present work uses a finite set of operational parameters under controlled laboratory conditions for inducing wear mechanisms and detecting them. However, in natural systems, the operational parameters vary dynamically. Thus, the robustness of the proposed framework should be tested, which is part of future work. Furthermore, with the ability to forecast the failure regimes with this methodology, the evolution of the reconstruction error can be used as a trigger to perform real-time imaging to gain additional insights into the wear mechanisms and understand their evolution; this is another planned research direction.

Acknowledgements

This work was funded by the Austrian COMET Program (project InTribology, No. 872176) via the Austrian Research Promotion Agency (FFG) and the Provinces of Niederösterreich and Vorarlberg, and has been carried out within the Austrian Excellence Centre of Tribology (AC2T research GmbH). The authors would like to thank Christoph Haslehner for performing the experiments and Matthias Freisinger for microscopic analysis of the bearings.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

[1] Mobley R K. *An Introduction to Predictive Maintenance*. Woburn (USA): Elsevier Inc., 2002.
[2] Swanson L. Linking maintenance strategies to performance. *Int J Prod Econ* **70**(3): 237–244 (2001)
[3] Durocher D B, Feldmeier G R. Predictive versus preventive maintenance. *IEEE Ind Appl Mag* **10**(5): 12–21 (2004)
[4] Carnero M C. Selection of diagnostic techniques and instrumentation in a predictive maintenance program. A case study. *Decis Support Syst* **38**(4): 539–555 (2005)
[5] McKone K E, Weiss E N. Guidelines for implementing predictive maintenance. *Prod Oper Manag* **11**(2): 109–124 (2002)
[6] Daily J, Peterson J. Predictive maintenance: How big data analysis can improve maintenance. In: *Supply Chain Integration Challenges in Commercial Aerospace*. Richter K, Walther J, Eds. Cham (Switzerland): Springer, Cham, 2017: 267–278.
[7] Selcuk S. Predictive maintenance, its implementation and latest trends. *Proc Inst Mech Eng B: J Eng Manuf* **231**(9): 1670–1679 (2017)
[8] Lu B, Durocher D B, Stemper P. Predictive maintenance techniques. *IEEE Ind Appl Mag* **15**(6): 52–60 (2009)
[9] Meng Y G, Xu J, Jin Z M, Prakash B, Hu Y Z. A review of recent advances in tribology. *Friction* **8**(2): 221–300 (2020)
[10] Evans D C. Self-lubricating bearings. *Ind Lubr Tribol* **33**(4): 132–138 (1981)
[11] Gawarkiewicz R, Wasilczuk M. Wear measurements of self-lubricating bearing materials in small oscillatory movement. *Wear* **263**(1–6): 458–462 (2007)
[12] Ren Y L, Zhang L, Xie G X, Li Z B, Chen H, Gong H J, Xu W H, Guo D, Luo J B. A review on tribology of polymer composite coatings. *Friction* **9**(3): 429–470 (2021)
[13] Paxton R R. *Manufactured Carbon: A Self-Lubricating Material for Mechanical Devices*. Boca Raton (USA): CRC Press, 2017.
[14] Bhushan B. *Modern Tribology Handbook, Two Volume Set*. Boca Raton (USA): CRC Press, 2000.
[15] Duan C J, He R, Li S, Shao M C, Yang R, Tao L M, Wang C, Yuan P, Wang T M, Wang Q H. Exploring the friction and wear behaviors of Ag–Mo hybrid modified thermosetting polyimide composites at high temperature. Friction 8(5): 893–904 (2020)

[16] Lancaster J K. Composite self-lubricating bearing materials. Proc Inst Mech Eng 182(1): 33–54 (1967)

[17] Xiang D H, Shan K L. Friction and wear behavior of self-lubricating and heavily loaded metal–PTFE composites. Wear 260(9–10): 1112–1118 (2006)

[18] Konstantinos. K. Tribology and condition monitoring of composite bearing liners for intelligent aerospace bearings. Ph.D. Thesis. Cardiff: Cardiff University, 2018.

[19] Deshpande P, Pandiyan V, Meylan B, Wasmer K. Acoustic emission and machine learning based classification of wear generated using a pin-on-disc tribometer equipped with a digital holographic microscope. Wear 476: 203622 (2021)

[20] Meylan B, Dogan P, Sage D, Wasmer K. A simple, fast and low-cost method for in situ monitoring of topographical changes and wear rate of a complex tribo-system under mixed lubrication. Wear 364–365: 22–30 (2016)

[21] Markus V, Reinhard G, Alexander M, Martin K. Online wear measurement in harsh environment. Part 1: Possible measurement strategies. Tribologie und Schmierungstechnik 66(4–5): 28–34 (2019) (in German)

[22] Markus V, Reinhard G, Alexander M, Martin K. Online wear measurement in harsh environment. Part 2: Application roller press. Tribologie und Schmierungstechnik 66(4–5): 35–43 (2019) (in German)

[23] Vakharia V, Gupta V K, Kankar P K. Ball bearing fault diagnosis using supervised and unsupervised machine learning methods. Int J Acoust Vib 20(4): 244–250 (2015)

[24] Orhan S, Aktürk N, Çelik V. Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool: Comprehensive case studies. NDT E Int 39(4): 293–298 (2006)

[25] Prieto M D, Cirrincione G, Espinosa A G, Ortega J A, Henao H. Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. IEEE Trans Ind Electron 60(8): 3398–3407 (2013)

[26] Sadegh H, Mehdi A N, Mehdi A. Classification of acoustic emission signals generated from journal bearing at different lubrication conditions based on wavelet analysis in combination with artificial neural network and genetic algorithm. Tribol Int 95: 426–434 (2016)

[27] König F, Sous C, Ouaid Chaïb A, Jacobs G. Machine learning based anomaly detection and classification of acoustic emission events for wear monitoring in sliding bearing systems. Tribol Int 155: 106811 (2021)

[28] Elforjani M, Shanbr S. Prognosis of bearing acoustic emission signals using supervised machine learning. IEEE Trans Ind Electron 65(7): 5864–5871 (2018)

[29] Glowacz A, Tadeusiewicz R, Legutko S, Caesarendra W, Irfan M, Liu H, Brunercik F, Gutten M, Sulowicz M, Antonino Daviu J A, et al. Fault diagnosis of angle grinders and electric impact drills using acoustic signals. Appl Acoust 179: 108070 (2021)

[30] Moder J, Bergmann P, Grün F. Lubrication regime classification of hydrodynamic journal bearings by machine learning using torque data. Lubricants 6(4): 108 (2018)

[31] Prost J, Cihak-Bayr U, Neacșu I A, Grundtner R, Pirker F, Vorlaufer G. Semi-supervised classification of the state of operation in self-lubricating journal bearings using a random forest classifier. Lubricants 9(5): 50 (2021)

[32] Mokhtari N, Pelham J G, Nowosisky S, Bote-Garcia J L, Gühmann C. Friction and wear monitoring methods for journal bearings of geared turbofans based on acoustic emission signals and machine learning. Lubricants 8(3): 29 (2020)

[33] Kankar P K, Sharma S C, Harsha S P. Fault diagnosis of ball bearings using machine learning methods. Expert Syst Appl 38(3): 1876–1886 (2011)

[34] Caesarendra W, Tjahjowidodo T. A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing. Machines 5(4): 21 (2017)

[35] Zhao B, Zhang X M, Zhan Z H, Wu Q Q. A robust construction of normalized CNN for online intelligent condition monitoring of rolling bearings considering variable working conditions and sources. Measurement 174: 108973 (2021)

[36] Eren L, Ince T, Kiranayaz S. A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier. J Signal Process Syst 91(2): 179–189 (2019)

[37] Wang D C, Guo Q W, Song Y, Gao S Y, Li Y B. Application of multiscale learning neural network based on CNN in bearing fault diagnosis. J Signal Process Syst 91(10): 1205–1217 (2019)

[38] Narendiranath Babu T, Aravind A, Rakesh A, Jahzan M, Prabha R D, Ramalinga Viswanathan M. Automatic fault classification for journal bearings using ANN and DNN. Archives of Acoustics 43(4): 727–738 (2018)

[39] Glowacz A. Ventilation diagnosis of angle grinder using thermal imaging. Sensors 21(8): 2853 (2021)

[40] Glowacz A. Fault diagnosis of electric impact drills using thermal imaging. Measurement 171: 108815 (2021)

[41] Lee K, Kim J K, Kim J, Hur K, Kim H. CNN and GRU combination scheme for bearing anomaly detection in rotating machinery health monitoring. In: Proceedings of the 1st IEEE International Conference on Knowledge Innovation and Invention, Jeju, Republic of Korea, 2018: 102–105.
Vigneashwara PANDIYAN. He is currently a postdoctoral researcher at the Laboratory for Advanced Materials Processing (LAMP), ETH Swiss Federal Laboratories for Materials Science and Technology (Empa, Switzerland). He completed his M.S. and Ph.D. from Nanyang Technological University (NTU), Singapore, under Rolls-Royce@NTU corporate lab. Prior to joining Empa, he was a research scientist in A*Star-Agency for Science, Technology and Research, Singapore. He now concentrates on implementing machine learning models for in-situ monitoring of manufacturing processes for anomaly detection based on sensor signatures.

Mehdi AKEDDAR. He is currently a master student at RELab at ETHZ on unsupervised home therapy for post stroke patients. He received a B.S. degree in microengineering from EPFL. He completed a master in medical robotics at École polytechnique fédérale de Lausanne (EPFL). He did a year-long visit to McGill, Montreal, Canada, in 2019 for the end of his B.S. degree. He completed an internship at Empa in the domain of machine learning applied to time series data from tribology.
Josef PROST. He received his Ph.D. in physics from Vienna University of Technology, Austria, in 2018. He is currently working as a postdoctoral researcher at the Austrian Excellence Centre for Tribology (AC2T research GmbH) in Wiener Neustadt, Austria. His main research interest is the application of advanced data analysis and visualisation methods to tribological research questions, including the detection of anomalous operation states and impending failures using machine learning models.

Georg VORLAUFER. He is currently a principal scientist at AC2T research GmbH. He completed his M.S. in physics at the TU Wien, Austria, in 1998 and received his Ph.D. degree in physics in 2002 from the same institution. Between 1998 and 2001, he carried out his Ph.D. studies in the field of vacuum and surface science at CERN (Geneva, Switzerland). He has more than 18 years of experience in the field of tribology. Although for many years his research interests have been mainly in the field of physics-based modelling and simulation of tribological systems, he is currently concentrating on tribology-related aspects of data science, machine learning, and artificial intelligence.

Markus VARGA. He is currently leading the strategic research area of Synaptic Tribology at AC2T research GmbH. He received his M.S. at the University of Applied Science Wiener Neustadt, Austria, in mechatronics and completed his Ph.D. in tribology at the Montanuniversität Leoben, Austria. For more than ten years, his main research field is the optimisation of industrial maintenance by tribological measures, i.e., wear protection, sensors for early detection of failures, etc.

Kilian WASMER. He received the B.S. degree in mechanical engineering from the Applied University, Sion, Switzerland and Applied University, Paderborn, Germany, in 1999. He received his Ph.D. degree in mechanical engineering from Imperial College London, UK, in 2003. His current position is deputy laboratory head of LAMP at Empa as well as a lecturer at EPFL.