INTERNATIONAL JOURNAL OF INTEGRATED ENGINEERING VOL. 12 NO. 3 (2020) 178-186

© Universiti Tun Hussein Onn Malaysia Publisher's Office

IJIE

Journal homepage: http://penerbit.uthm.edu.my/ojs/index.php/ijie
ISSN : 2229-838X e-ISSN : 2600-7916

The International Journal of Integrated Engineering

A Deep Learning Approach to Prognostics of Rolling Element Bearings

Jang-Wook Hur, Ugochukwu Ejike Akpudo*

Kumoh National Institute of Technology, 61 Daehak-ro (yangho-dong), Gumi, Gyeongbuk, 39177, KOREA

*Corresponding Author

DOI: https://doi.org/10.30880/ijie.2020.12.03.021
Received 18 December 2019; Accepted 14 January 2020; Available online 27 February 2020

Abstract: The use of deep learning approaches for prognostics and remaining useful life predictions have become obviously prevalent. Artificial recurrent neural networks like the long short-term memory are popularly employed for forecasting, prognostics and health management practices, and in other fields of life. As an unsupervised learning approach, the efficiency of the long short-term memory for time-series predictive purposes is quite remarkable in contrast to standard feedforward neural networks. Virtually all mechanical systems consist mostly of rotating components which are by nature, prone to degradation/failure from known and uncertain causes. As a result, condition monitoring of these rolling element bearings is necessary in order to carry out prognostics and make necessary life predictions which guide safe and cost-effective decision making. Several studies have been conducted on effective approaches and methods for accurate prognostics of rolling element bearings; however, this paper presents a case study on rolling element bearing prognostics and degradation performance using an LSTM model.

Keywords: Bearing Degradation, Long short-term memory, Feature Extraction, Prognostics, Degradation assessment.

1. Introduction

Prognostics and health management approaches have witnessed a global improvement ranging from the use of traditional physics-based models to more robust data-driven models [1]. The necessity for condition monitoring and prognostics of rotating components like bearings cannot be overemphasized [2]. Not only does it offer a wider range of information for decision-making, but also, the information gathered from prognostics inspire predictive maintenance decisions for extended lives of systems, cost optimization, safety, and productivity.

The use of deep learning approaches for prognostics and remaining useful life (RUL) predictions have been on the increase and research studies are constantly being conducted for more effective and optimized approaches. Statistical machine learning models, the more robust deep learning models, and a host of hybrid models are being used and customized for optimum prognostics needs in various areas of life. These data-driven models are designed to accept a feature vector as input and make predictions as output [3]. Artificial recurrent neural networks like the long short-term memory (LSTM) are recently being popularly employed for prognostics and health management (PHM) purposes and in other areas of life [4]-[6]. As an unsupervised learning approach, the LSTM's efficiencies for time-series predictive purposes are quite remarkable in contrast to standard feedforward neural networks (FNNs) whose architecture has only a forward process- from the input nodes, through the hidden layers and finally to the output nodes [7].

Several recent research studies have been conducted on effective and cost-efficient approaches to data-driven prognostics based on recurrent neural network (RNN) architecture [8]-[9]; however, because a typical LSTM architecture comprises of input-hidden-output layers with a feedback structure, it uses gates to control the memorizing process [10]. This addresses the challenging problems of vanishing and exploding gradient issues which conventional RNNs are prone to; consequently, making them one of the 'general purpose computers' [11].

Corresponding author: ugoakpudo@kumoh.ac.kr

Virtually all mechanical systems consist mostly of rotating components which are naturally prone to failure from both known, and uncertain causes. This is because rolling element bearings (REBs) are commonly used in mechanical systems; therefore, condition monitoring (CM) of these REBs is necessary in order to carry out prognostics and make necessary life predictions which guide safe, and cost-effective decision making. Accurate health assessment is by a norm, a primary pre-requisite for enabling predictive maintenance procedure for increased productivity and reliability as well as ensuring optimized utility of machine components. Many works of literature have presented analytical models for asset health assessment which relied primarily on historical operation data and sensor data. The authors of [12] presented a health assessment approach for REBs by employing the Hilbert-Huang transform as a representative degradation feature as input for a particle filter model for tracking degradation state of the test bearings. Consequently, a posterior probability distribution of RUL was obtained. From [13], the authors developed a Paris law model comprising of a physical-based and a statistical-based approach to RUL prediction of REBs. From the health assessment model of rotating bearing proposed by [14], an empirical mode decomposition method was employed for feature extraction from raw vibration signals followed by a self-organizing map method for deducing the bearing's health state based on features that were extracted.

As seen from the above studies (and in many others works of literature), the general ideology for implementing data-driven prognostics revolves around the processes: feature extraction for showing system behaviour, construction of a prediction model for assessing current system health, and for making future predictions; and finally, RUL prediction. In view of this, this study presents a case study of REB prognostics and degradation performance using an LSTM model. First, feature extraction was carried out on the test bearings, followed by a degradation assessment of the bearings. Using the LSTM model, future predictions were made using the extracted feature.

The remaining part of the paper is structured as follows: Section II presents a review of the LSTM. Section III shows a degradation assessment model based on the LSTM model. Section IV presents an experimental study of the proposed model using REB data provided by Intelligent Maintenance System (IMS) Center, University of Cincinnati [15]. Section V concludes the paper.

2. Overview of the Long Short Term Memory

The versatility in purpose of the LSTM ranging from speech recognition [16], language modelling, text prediction [17], etc. cannot be overemphasized owing to its feedback-integrated architecture, unlike other RNNS.

In a typical LSTM network, an LSTM cell contains three gates, namely the input gate, forget gate, and output gate [18]. These gates control information passage along with the sequences which can acquire long-range dependencies with more accuracy [19]. Fig. 1a shows a typical LSTM cell architecture while fig. 1b shows a series of LSTM cells at time steps.

The forget gate , input gate I_t , and output gate O_t are single layered neural networks which contain sigmoid activation function (yields output between 0 and 1) while the candidate layer uses the $tanh$ activation function (yields output between -1 and 1). These gates take the input vectors(u), and previous output vectors (w), concatenate them, and finally apply the sigmoid activation function.

From fig. 1a, the first layer $\boldsymbol{f_t}$ determines what information to transfer from the previous state $\boldsymbol{C_{t-1}}$ and considers a previous output $\boldsymbol{H_{t-1}}$ and current input $\boldsymbol{x_t}$. Mathematically, it is expressed in (1) as:

$$= (\boldsymbol{u_f x_t} + \boldsymbol{w_f H_{t-1}} + \boldsymbol{b_f}) \quad (1)$$

The second layer (i_t) decides what information to be stored in the current state and is shown in (2).

$$= (\boldsymbol{u_i x_t} + \boldsymbol{w_i H_{t-1}} + \boldsymbol{b_i}) \quad (2)$$

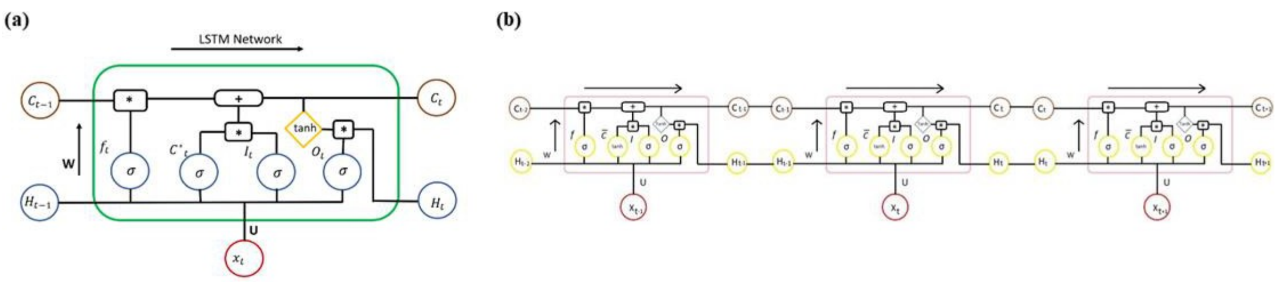

Fig. 1 - (a) LSTM architecture showing a single LSTM cell; (b) LSTM architecture showing LSTM cells at different time steps

Where σ represents the sigmoid activation function and w_i represents the weight of f_t. w_i represents the weight of layer i while b_i represents the bias of layer i.

The current cell memory C_t shown in (3) is a function of the previous memory $C_{t\text{-}1}$ and emanates from an element-wise multiplication with the forget gate f. The updated candidate from the previous cell to current cell memory is expressed mathematically in (4) as:

$$C^*{}_t \;=\; h(w_c \circledast H_{t\text{--}1} + u_c x_t + b_c) \tag{3}$$

$$C_t = \;\circledast\, C_{t-1} \oplus i_t \cdot C^*{}_t \tag{4}$$

Each LSTM cell output H_t is a function of yet another element-wise multiplication between the $tanh$ activated function of C_t, and the last layer- output layer O_t and are shown in (5) and (6) respectively.

$$= (\boldsymbol{u_o x} + \boldsymbol{w_o H_{t-1}}) + \boldsymbol{b_0}) \tag{5}$$

$$= \;\circledast\, \mathbf{tanh}(\boldsymbol{C_t}) \tag{6}$$

Eventually, the first hidden layer outputs ($\boldsymbol{C_t}$ and $\boldsymbol{H_t}$) becomes the input layer of the next cell to form a neural network of several hidden LSTM layers (or cells). Table 1 describes each parameter.

Table 1 - Description of Parameters

Symbol	Description
x_t	Input Vector
H_{t-1}	Hidden Previous cell output
C_{t-1}	Previous cell memory
H_t	Current cell output
C_t	Current cell memory
$\circledast$	Element wise multiplication
$\oplus$	Element-wise addition
w, u	Weight vectors for f_t, $C^*{}_t$, I_t, O_t
f_t	Forget gate
$C^*{}_t$	Candidate
I_t	Input gate
O_t	Output gate

In summary, the goal is to capture prior information that may contribute to the posterior event in a feedback-enabled manner. This leads to the output which is an averaged sum over time and is represented in (7) as:

$$\boldsymbol{H} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{H_j / n} \tag{7}$$

3. Degradation Assessment Model

This section presents the procedure for carrying out degradation assessment using the proposed method. These include feature extraction and an overview of the proposed prognostics model

3.1 Feature Extraction

In order to effectively carry out effective degradation assessment, the primary and most sensitive step is to extract feature(s) that show the degradation trend from normal health condition until failure state. Fig. 2 shows the raw vibration signal for the data points however, from the plot, one cannot make meaningful deductive conclusions.

Vibration signals have shown in [20], [21] and many other works of literature to reflect the behavior of rotating element bearings and from these signals, frequency-domain features (Hilbert–Huang transform, fast Fourier transform, etc.), time-domain features (root mean square, Kurtosis, etc.) or time-frequency-domain features (Short-time Fourier transform, Wigner transform, etc.) can be extracted and prepared for prognostics.

As verified in [22], the RMS, as a single degradation feature is an effective time-domain statistical feature for representing degradation behavior of vibrational (and other signals) with relatively lesser noise (compared to other time-

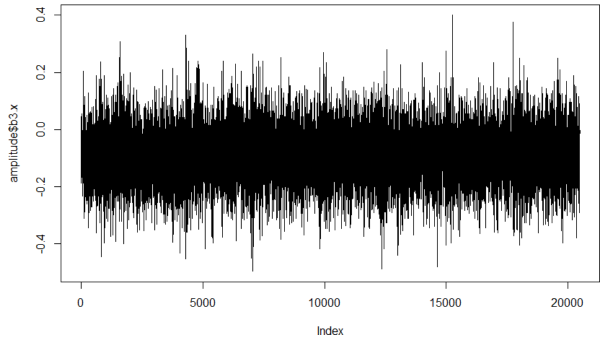

Fig. 2 - Raw vibration signal of Bearing4

based features) and can successfully extract valuable information from the vibrational signals. This study presents the RMS as a key feature for assessing bearing degradation behavior and is represented in (8) as:

$$X_{RMS} = \sqrt{\frac{\sum_{k=1}^{k}((k))^{2}}{k}} \qquad (8)$$

Where x(k) is a series of vibration signal for k = 1, 2, 3, …, K.

3.2 Proposed Prognostics Model

Fig. 3 shows the schematic prediction model for the proposed study. After feature extraction and normalization, the normalized features were split into training and test data. The training data was input to the model for, and upon completion, was fed with the test data for prediction.

As suggested by the authors of [23], the training-testing process was carried out using several combinations of parameters like the epoch size, batch number, number of LSTM cells. This continued iteratively (by trial and error) until an acceptable model was achieved. By observing the loss function of every model using several combinations of above-mentioned parameters, the model with best convergence of the loss function to zero, is most likely a befitting model for prediction; however, the model's predictive accuracy chiefly depends its performance on the test data.

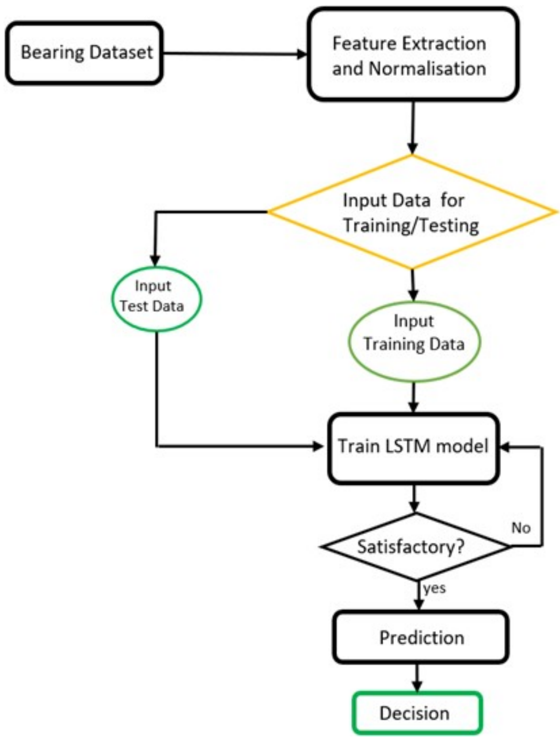

Fig. 3 - Overview of LSTM model for bearing prognostics

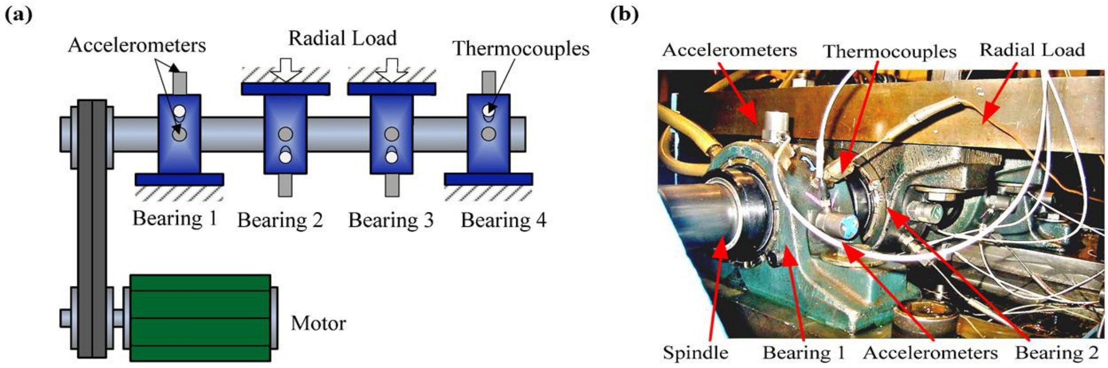

Fig. 4 - (a) Schematic view showing Sensor placement and loading illustration; (b) Actual picture of test bench (see [24])

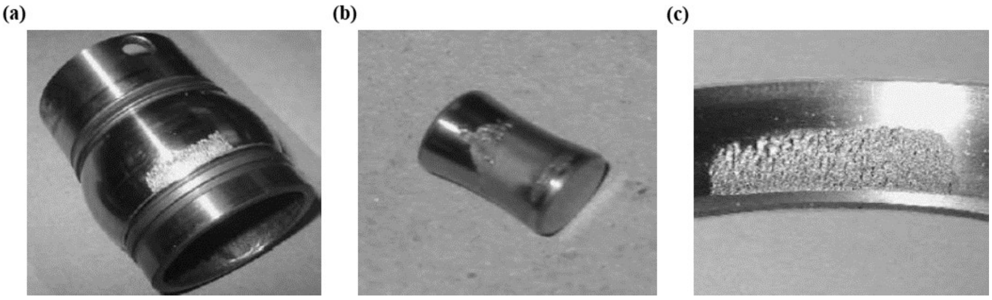

Fig. 5 - (a)Inner race failure in bearing3; (b) roller element failure in bearing4; (c) outer race failure in bearing1 (see [24])

4. Analysis Using The Proposed Method

In this section, the paper presents the source of vibrational data and the experimental results from the analyses.

4.1 Experimental Setup

The REB data was generated from a run-to-failure test of four bearings fitted on a loaded shaft with a radial load of 2730kg. Data was collected hourly with a sampling frequency of 2000rpm. Through a NI DAQ Card-6062E, the vibrational data was collected via vertically (and horizontally) installed high sensitivity accelerometers. Three run-to-failure experiments were carried out and the results were recorded in three datasets. The data was provided by Intelligent Maintenance System (IMS) Centre, University of Cincinnati [15]. As originally presented in [24, fig. 16], fig. 4 shows the experimental test setup with a schematic view showing Sensor placement and loading illustration.

4.2 Experimental Results

Fig. 5 shows the results from the run-to-failure test on the bearings. The results showed an inner race failure in bearing3 and a rolling element failure in bearing4, and an outer race failure in bearing1.

4.3 Feature Extraction

The RMS feature of the failed bearings were extracted, normalized (using a moving average algorithm of window size 10), and presented in figs. (6-8). As seen in Fig. 6, Bearing3 failed uniformly as seen from the degradation pattern; however, bearing4 exhibited a more random behaviour than bearing3 (Fig. 7). On the other side, bearing1 showed an early suspicious behaviour but later showed some self-healing effects as seen in fig. 8.

Owing to the anomaly in failure behaviour of bearing4 towards its end of life (bearing4 experienced a stage 2 failure before finally experiencing a rolling element failure), this study shall focus more on vibrational signal of bearing4 for analysis. As seen from Fig. 7, it would take an effective prediction model to make accurate predictions from the 1600th measurement point, onwards. We aim to show the dynamically robust capability of the proposed model in making accurate forecasts and future predictions from the test data after training. This shall further be illustrated in section D.

At about 1750 measurement point (corresponding to day 26), the RMS exceeded the failure threshold (0.18); hence, signaling failure. Nevertheless, the bearing showed some sort of self-healing behaviour but was short-lived from about day 28 till it finally failed (stage 2 failure).

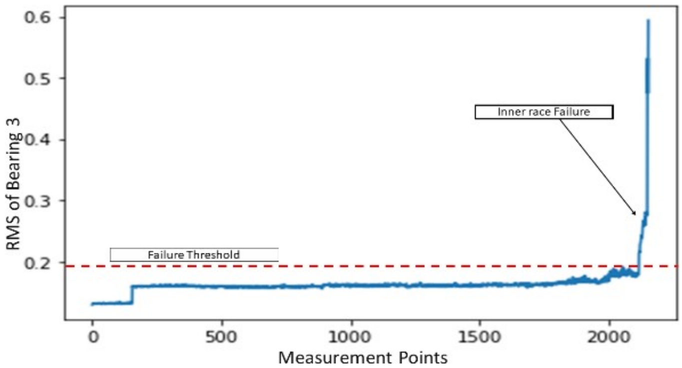

Fig. 6 - RMS of bearing3

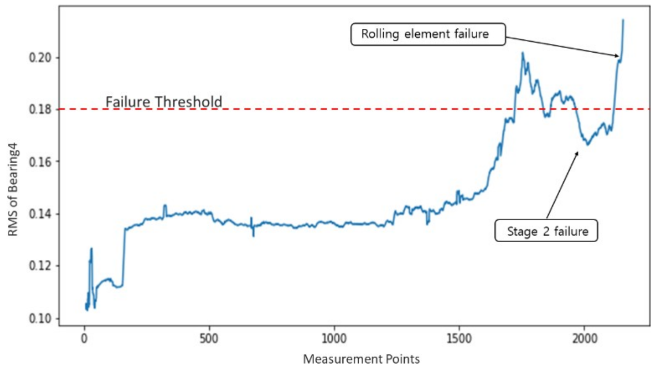

Fig. 7 - RMS of bearing4

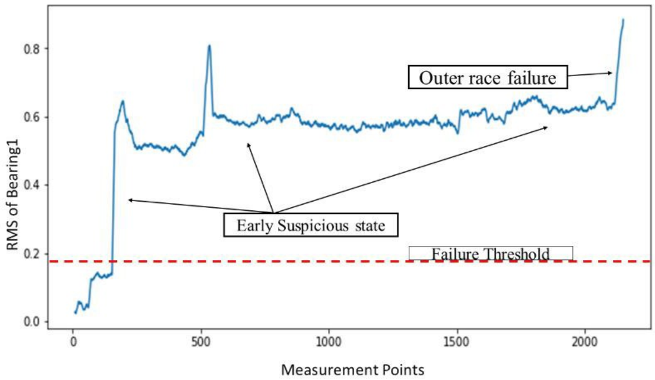

Fig. 8 - RMS of bearing1

4.4 Degradation assessment based on LSTM model

The accuracy in the prediction of an LSTM model depends reasonably on the number of hidden layers in the network. By stacking more LSTM cells or hidden layers, a deep LSTM network with increased model complexity and predictive power can be built [16]. More empirically, parameters like learning rate, epoch size, number of time-step-sized batches available to be iterated in each epoch, how densely the neurons in the network are connected, etc., all determine the performance of an LSTM algorithm on an input data.

After feature extraction and normalization of bearing4 vibrational signal, the normalized data was divided into training, and test data in the ratio 70:30. The training data was used to train the proposed LSTM model, and after training, was fed with the test data to make predictions. Fig 9 shows the convergence of the loss function of the trained model. For the proposed model, several trials using several combinations of earlier mentioned parameters were conducted. These combinations are recorded in Table 2 and their corresponding results presented in fig. 10. After several trials with various parameters, fig. 10f shows the most accurate predictive performance of the proposed model on bearing4 RMS against other recorded parameter combinations. This corresponds to an LSTM models of 5 hidden layers with epoch size of 50 and a batch size of 40.

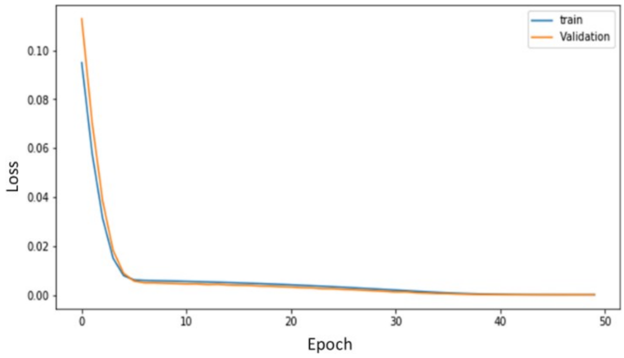

Fig. 9 - Loss function of the trained model

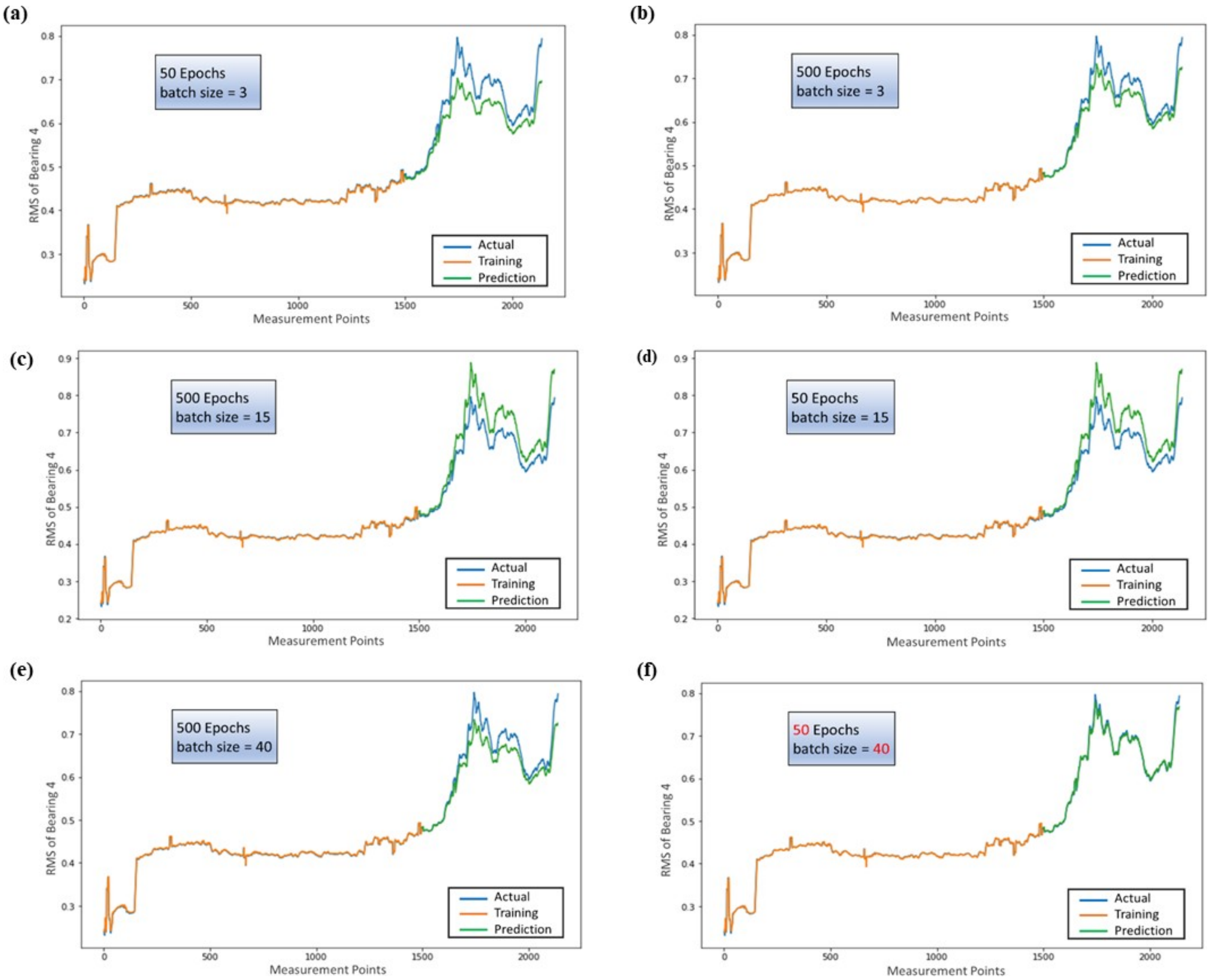

Fig. 10 - RMS degradation predictions of bearing4 at various combinations of Epochs and batch sizes

Table 2 - RMSE Comparison of Model Parameters

Epoch size	Batch Size	No. of LSTM cells	Computation time(secs)	RMSE (%)
50	3	5	88	2.76
500	3	5	213	2.31
500	15	5	157	2.56
50	15	5	65	2.53
500	40	5	129	1.75
50	40	5	41	1.1

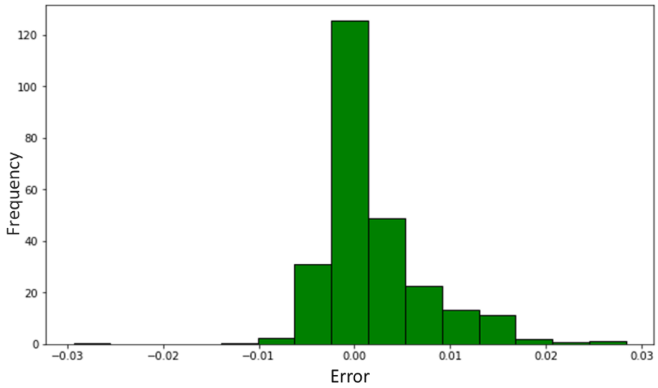

Fig. 11 - Prediction error histogram of proposed model on bearing4

Table 3 - Prediction Results of Proposed Model

Performance Metric	Bearing1	Bearing3	Bearing4
RMSE	1.4%	1.2%	1.1%

4.5 Performance Evaluation Criteria

The RMS as seen from the previous section represented effectively, the degradation behavior of the bearing. The Root Mean Square Error (RMSE) is employed for evaluating the proposed model's accuracy in predicting the bearing's degradation behaviour and is expressed in (9) as:

$$\text{RMSE} \quad = \sqrt{\frac{1}{n}\sum\left(x(i) - x^{*}(i)\right)^{2}} \qquad (9)$$

Where (i) is the i-th value of bearing RMS, $x^{*}(i)$ represents the i-th RMS prediction value.

4.6 Evaluation Result

Table 3 presents the prediction results of the proposed LSTM model on bearing1, bearing3 and bearing4. Fig. 11 shows the RMSE distribution of the proposed model. As seen from the results in table III, the proposed model has a 99.9% average RMS prediction accuracy on the presented REBs.

5. Conclusion and Future Works

The proposed approach is based on a deep learning architecture for predicting bearing degradation behavior using a run-to-failure data provided by the NASA. The deep learning structure comprises layers of LSTM neurons and was modeled using data from the run-to-failure experiment. As a prerequisite, feature extraction was carried out, and as a result, RMS values for the whole degradation life of the REBs were obtained, normalized and for degradation assessment.

The proposed model was validated using the RMS values extracted from the dataset (training data), and with the test data, degradation prediction was made by the trained LSTM model. Consequently, the results show that the proposed model can capture system dynamic behavior, and detecting fault in REBs with admirable accuracy, and with less effort, unlike most traditional physics-based prognostics models.

In further studies, we shall carry out a more comprehensive feature extraction from the test bearings to really capture in full, the bearing degradation behavior. As suggested by the authors of [25], by combining other time-based features with the RMS, a more comprehensive feature index can be obtained and can serve as a better health indicator.

In addition, we shall develop a more effective hybrid model for prognostics by combining LSTM and physics-based prognostics methods. This shall ensure that not only will the physical behavior of the bearing be captured, but no stone would also be left unturned in the whole process of data analysis, feature extraction, degradation assessment, and prognostics. The new model shall not only be implemented on bearings, but also, on other rotating components like fluid pumps, brushless DC Motors, etc.

Acknowledgement

This article is based on the results of a study conducted with the support of the Agency for Defense Development (RAM specialized laboratory, UD180018AD).

References

[1] Cahyo, W. N., Prawahandaru, H., Swasono, B. A., Raben, R. S. I., Sutartono, R. T., & Immawan, T. (2019). Data-Based Maintenance Strategy Analysis using Operational Excellence Approach in Engineering Asset Management. *International Journal of Integrated Engineering*, 11(5), 222-228.

[2] Mohd Idrus, M. N. E., Chia, K. S., Sim, H. M., & Gamal Al-kaf, H. A. (2018). Artificial Neural Network and Savitzky Golay Derivative in Predicting Blood Hemoglobin Using Near-Infrared Spectrum. *International Journal of Integrated Engineering,* 10(8), 181-193

[3] L. C. Jaw. (1999). Neural Networks For Model-Based Prognostics. *1999 IEEE Aerospace Conference Proceedings (Cat. No.99TH8403),* Snowmass At Aspen, CO, USA. Vol.3, 21-28.

[4] Mohd Idrus, M. N. E., Chia, K. S., Sim, H. M., & Gamal Al-kaf, H. A. (2018). Artificial Neural Network and Savitzky Golay Derivative in Predicting Blood Hemoglobin Using Near-Infrared Spectrum. *International Journal of Integrated Engineering,* 10(8), 112-119.

[5] M. Madkour, D., Ahmed, M., & Mohamed, W. F. (2019). Automatic Face and Hijab Segmentation Using Convolutional Network. *International Journal of Integrated Engineering*, 11(7), 61-66.

[6] Md Yunus, M. A. (2017). Artificial Neural Network and Wavelet Features Extraction Applications in Nitrate and Sulphate Water Contamination Estimation. *International Journal of Integrated Engineering*, 9(4), 64-75

[7] Zhang, Yongzhi Et Al. (2017). A LSTM-RNN Method For The Lithuim-Ion Battery Remaining Useful Life Prediction. Prognostics And System Health Management Conference (PHM-Harbin), 1-4.

[8] Renato Giorgiani, Nascimento, Felipe A. C. Viana. (2019). Fleet Prognosis With Physics-Informed Recurrent Neural Networks. *Arxiv:1901.05512v1 [Cs.CE].*

[9] Linxia Liao, Hyung-Il Ahn. (2016). Combining Deep Learning And Survival Analysis For Asset Health Management. *International Journal Of Prognostics And Health Management*, 7(020):7

[10] Siegelmann, Hava T., Sontag, Eduardo D. (1992). On The Computational Power Of Neural Nets. *ACM. COLT'92,* 440–449.

[11] X. Liu, P. Song, C. Yang, C. Hao And W. Peng. (2018). Prognostics And Health Management Of Bearings Based On Logarithmic Linear Recursive Least-Squares And Recursive Maximum Likelihood Estimation. *IEEE Transactions On Industrial Electronics,* Vol. 65, No. 2, 1549-1558.

[12] M. Behzad, H. A. Arghan, A. R. Bastami And M. J. Zuo. (2017). Prognostics Of Rolling Element Bearings With The Combination Of Paris Law And Reliability Method. *Prognostics And System Health Management Conference (PHM-Harbin)*, Harbin, 1-6.

[13] Hong, S., Zhou, Z., Zio, E., &Wang,W. (2014). An Adaptive Method For Health Trend Prediction Of Rotating Bearings. *Digital Signal Processing*, 35, 117–123.

[14] J. Lee, H. Qiu, G. Yu, J. Lin, And Rexnord Technical Services (2007). IMS, University Of Cincinnati. "Bearing Data Set", *NASA Ames Prognostics Data Repository,* NASA Ames Research Center, Moffett Field, CA. Available at: http://Ti.Arc.Nasa.Gov/Project/Prognostic-Data-Repository

[15] Medennikov, Ivan & Bulusheva, Anna. (2016). LSTM-Based Language Models For Spontaneous Speech Recognition. *International Conference On Speech And Computer*, 469-475

[16] M. Sundermeyer, R. Schl¨Uter, And H. Ney. (2012). LSTM Neural Networks For Language Modeling. *13th Annual Conference Of The International Speech Communication Association, INTERSPEECH.*

[17] Wang, Y.; Xie, D.; Wang, X.; Zhang, Y. (2018). Prediction Of Wind Turbine-Grid Interaction Based On A Principal Component Analysis-Long Short Term Memory Model. *Energies*, 11, 3221

[18] Rui Zhao, Ruqiang Yan, Jinjiang Wang, Kezhi Mao. (2017). Learning To Monitor Machine Health With Convolutional Bi-Directional LSTM Networks. *Intelligent Sensing And Information Mining, Sensors*, 17(2), 273.

[19] Haidong Shao, Hongkai Jiang, Xingqiu Li, Tianchen Liang. (2018). Rolling Bearing Fault Detection Using Continuous Deep Belief Network With Locally Linear Embedding. *Computers In Industry*, vol. 96, 27-39.

[20] Yusuf, A. I., Mohd Amin, N., Yunus, M. A., & Abdul Rani, M. N. (2019). Evaluation of Rayleigh Damping Coefficients for Laminated Rubber Bearing Components Using Finite Element and Experimental Modal Analysis. *International Journal of Integrated Engineering*, 10(9), 17-22

[21] Fatih Camci, Kamal Medjaher, Noureddine Zerhouni, Patrick Nectoux. (2012). Feature Evaluation For Effective Bearing Prognostics. *Quality And Reliability Engineering International*, Wiley, 1-15.

[22] Ian Goodfellow, Yoshua Bengio, Aaron Courville. (2016). Adaptive Computation And Machine Learning Series-Deep Learning-*The MIT Press*. P. 429.

[23] Hai Qiu, Jay Lee, Jing Lin. (2006). Wavelet Filter-Based Weak Signature Detection Method And Its Application On Roller Bearing Prognostics. *Journal Of Sound And Vibration*, vol. 289, 1066-1090.

[24] Vepa Atamuradov, Fatih Camci. (2017). Segmentation Based Feature Evaluation And Fusion For Prognostics. *International Journal Of Prognostics And Health Management,* vol. 8 (ISSN 2153-2648).