

A DEEP LONG SHORT-TERM MEMORY NETWORK FOR BEARING FAULT DIAGNOSIS UNDER TIME-VARYING CONDITIONS

Kai Zhou

Department of Mechanical Engineering-Engineering
Mechanics
Michigan Technological University
Houghton, MI 49931
kzhou@mtu.edu

ABSTRACT

The fault diagnosis of bearing in machinery system plays a vital role in ensuring the normal operating performance of system. Machine learning-based fault diagnosis using vibration measurement recently has become a prevailing approach, which aims at identifying the fault through exploring the correlation between the measurement and respective fault. Nevertheless, such correlation will become very complex for the practical scenario where the system is operated under time-varying conditions. To fulfill the reliable bearing fault diagnosis under time-varying condition, this study presents a tailored deep learning model, so called deep long short-term memory (LSTM) network. By fully exploiting the strength of this model in characterizing the temporal dependence of time-series vibration measurement, the negative consequence of time-varying conditions can be minimized, thereby improving the diagnosis performance. The published bearing dataset with various time-varying operating speeds is utilized in case illustrations to validate the effectiveness of proposed methodology.

Keywords: bearing fault diagnosis, time-varying conditions, time-series vibration measurement, deep long short-term memory (LSTM) neural network.

1. INTRODUCTION

Rolling bearing, as one key functional element in the rotating machinery has been extensively utilized. Due to the harsh and long-time operating conditions, bearing is prone to be failure. To maintain the bearing health and thus ensure the normal system function, fault diagnosis and prognosis of bearing recently have become a critically important subject. It is well known that the vibration signals directly measured from the machinery system are commonly used to facilitate the fault diagnosis owing to their notable advantages over other types of signals [1–3]. By further taking advantage of signal processing analysis, the fault-related features in the vibration measurement

will be extracted [4–6]. While those features appear to be the effective indicator of faults, discriminating features manually to pinpoint the faults in many scenarios is difficult and unreliable.

With the advancement of the computational intelligence technology, the machine learning methods have been substantially applied in fault diagnosis because of its capability in characterizing the inherent correlation between the measurement and faults [7–9]. Deep learning, as one representative class of machine learning methods has become a prevailing approach. A large-scale neural network with large number of layers and nodes oftentimes is established to fulfill the deep learning analysis. To improve the feature extraction ability and efficiency, the convolutional layers are integrated into the neural network to form so called convolutional neural network (CNN), which has been employed in a wide range of studies. Chen et al [10] proposed the combination of CNN and cyclic spectral coherence (CSCoh) to improve the performance of bearing fault diagnosis. Zhu et al [11] developed a new CNN built upon the capsule network to address the lack of positional relationship between fault features in traditional CNN, thereby enhancing the generalization ability. Hoang and Kang [12] achieved the automatic fault diagnosis directly upon the raw vibration signals by utilizing the deep structure of CNN.

It is worth mentioning that, despite the success of deep learning-based fault diagnosis, the challenges still exist in actual practice. Recently, the extensive effort has been made to address the challenges and build gap between the academic research and industrial applications. For example, Ma et al [13] proposed transfer learning together with the time-frequency domain signal processing to improve the fault diagnosis reliability under small-sized dataset. Zhou and Tang [14] developed a new fault diagnosis framework built upon the fuzzy inference system to tackle the issue of limited fault labels which cannot represent all unseen fault scenarios in practical implementations. The objective of this research is to deal with another notable challenge caused by the time-varying conditions. Considering

the fact that the bearings often operate under time-varying rotational speed conditions, the vibration signals measured are non-stationary, which poses the difficulty in fault pattern recognition using deep learning analysis. In this research, built upon the CNN architecture, long short-term memory (LSTM) features are further integrated to establish a deep long short-term memory (DLSTM) network, which aims at alleviating the non-stationary issue through characterizing the temporal correlation of time-series vibration signals.

The rest of the paper is organized as follows. Section 2 outlines the DLSTM network where its particular features to handle the correlation among time sequences are highlighted. Section 3 implements the fault diagnosis practice on a public dataset using the proposed methodology. Concluding remarks are given in Section 4.

2. METHODOLOGY

The tremendous progresses on the deep learning neural network development have been achieved over the last decade. Different network architectures and learning strategies have been established to realize various purposes. Amongst them, LSTM network is particularly employed in this research. The unique advantage of LSTM network is its ability in elucidating the correlation between the fault and features by taking the temporal dependence of features into account, which is tailored for tackling the non-stationary issue induced by the time-varying conditions. For this reason, LSTM network has been adopted in a broad of engineering applications [15–17].

LSTM network essentially is a special variant of recurrent neural network (RNN). The unique component embedded in the LSTM network is the memory cell, which can memorize the temporal state by 3 different controlling gates including input, forget and output gates as shown in Figure 1a. The underlying principle of collaboration among those gates can be mathematically formulated as

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \quad (1a)$$

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \quad (1b)$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \quad (1c)$$

$$c_t = f_t \odot C_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \quad (1d)$$

$$m_t = o_t \odot h(C_t) \quad (1e)$$

where f_t, i_t , and o_t denote the forget, input and output gates, respectively. c_t and m_t denote the cell state and cell output, respectively. W and b are the weight matrices and bias vectors to be optimized through network training. $\odot$ denotes the scalar product. x_t represents t -time sequence in one sample. σ is the sigmoid activation function, and g and h denote the hyperbolic tangent activation functions. The output of the cell is expressed as

$$y_t = W_{ym}m_t + b_y \quad (2)$$

Multiple memory cells are combined to form a layer (Figure 1b), which is referred to as LSTM layer. It is able to characterize the temporal correlation of time-series signals. Since the time-series vibration signals are lengthy especially when they are

measured from the real-time in-service machinery system, the convolutional layers are necessarily integrated into the LSTM network to realize the effective feature extraction. To reduce the network complexity and thus improve the training efficiency, maxpooling layers oftentimes are placed after the convolutional layers for downsampling. Usually, the convolutional layers together with maxpooling layers that are placed at the beginning of network are intended for low-level feature extraction, whereas the fully connected layers after LSTM layer are responsible for building the mapping between the fault conditions and high-level features obtained. For each layer, specific activation function is adopted to incorporate the data nonlinearity.

With all necessary layers integrated, a deep LSTM (DLSTM) network will be established. Once it is adequately trained through backpropagation optimization, it can be readily applicable for bearing fault diagnosis in this research.

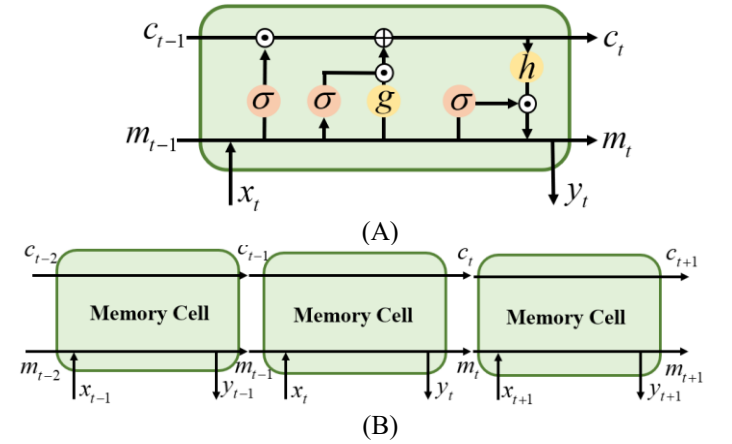


FIGURE 1. KEY ELEMENT IN DLSTM (A) MEMORY CELL; (B) LSTM LAYER.

3. CASE ILLUSTRATION

In this section we implement the methodology presented earlier to conduct comprehensive illustrations.

3.1 Experimental setup and problem formulation

To facilitate the methodology validation, a dataset that was reported in [18] is utilized in this research. The experimental testbed is shown in Figure 2. The experimental data acquisition was performed on a SpectraQuest machinery fault simulator. The experimental bearing placed at the right side of testbed is subject to three fault conditions, i.e., healthy condition, inner race fault and outer race fault. An encoder and an accelerometer were used to measure the shaft rotational speed and vibration signals, respectively. The shaft rotational speed was controlled to generate the multiple scenarios on purpose. The details of operating set-up are given in Table 4. Specifically, four speed varying tendencies were considered, each of which contains three different scenarios. Considering three fault conditions mentioned above, totally 36 (i.e., $3 \times 3 \times 4$) groups were created. The sampling frequency used was 20,000 Hz, and the time

duration for data acquisition was 10 s. With such set-up, the size of entire dataset can be represented as $36 \times 2,000,000$.

Generally, the time-series signals in all groups can be segmented into many samples to accommodate the subsequent deep learning analysis. In this research, each sample consists of 0.1 s time series with 20,000 data points. As such, 100 samples are generated for each group and totally 3,600 samples are involved in the entire dataset. It is worth nothing that, to highlight the advantage of the proposed methodology, the time-varying speed conditions of samples for testing are intended to be beyond that of samples for training. Considering that scenarios (a), (b) and (c) in each speed varying tendency (shown in Table 1) have certain level of similarity, only the data from scenario (c) are kept, which facilitates the fault diagnosis analysis under unseen/unexpected time-varying conditions. Hence, the original 3,600 samples are reduced to 1,200 samples for following analysis. The overview of the new reduced-sized dataset is given in Table 2. To thoroughly examine the fault diagnosis performance, four testing cases are formulated, where in each case the 200 samples collected upon certain speeds for both healthy and outer race fault are used for testing and the rest, i.e., 1,000 samples are used for training. We don't involve the inner race fault samples for testing because their vibration amplitudes generally are more significant than that of other fault types, as shown in Figure 3. In other words, the effect of time-varying speed may become negligible in differentiating the inner race fault from other faults, which is not beneficial for methodology validation.

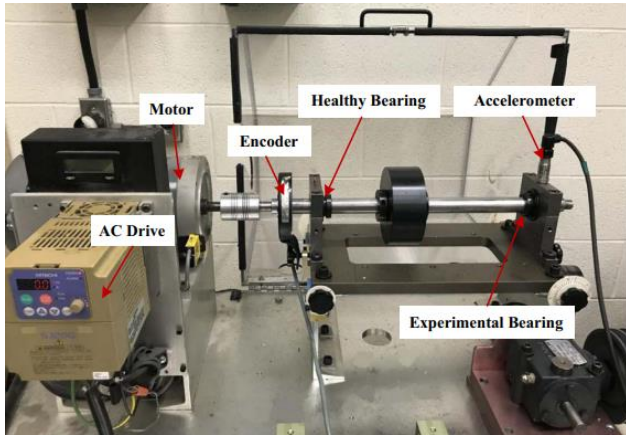


FIGURE 2. EXPERIMENTAL TESTBED [18]

Table 1. EXPERIMENTAL OPERATING SETUP [18]

Operating Rotational Speed (Hz)				
	Speed increasing	Speed decreasing	Speed first increasing, then decreasing	Speed first decreasing, then increasing
Healthy	(a). From 14.1 to 23.8	(a). From 28.9 to 13.7	(a). From 14.7 to 24.2	(a). From 24.2 to 14.7

Inner race fault	(b). From 14.1 to 29.0	(b). From 25.7 to 11.6	(b). From 25.3, then to 21.0 (b). From 14.4 to 24.0, then to 18.7	(b). From 14.8, then to 20.6 (b). From 24.6 to 14, then to 20.6
	(c). From 15.2 to 26.7	(c). From 28.6 to 13.9	(c). From 15.4 to 24.8, then to 19.1	(c). From 26 to 16.9, then to 23.2
	(a). From 12.5 to 27.8	(a). From 24.3 to 9.9	(a). From 15.1 to 24.4, then to 18.7	(a). From 25.3 to 14.8, then to 19.4
	(b). From 13.0 to 25.7	(b). From 25.1 to 13.1	(b). From 14.1 to 23.5, then to 18.0	(b). From 25.3 to 15.1, then to 19.8
	(c). From 13.5 to 28.5	(c). From 25.8 to 12.0	(c). From 14.8 to 21.7, then to 13.6	(c). From 23.1 to 15.7, then to 23.6
	(a). From 14.8 to 27.1	(a). From 24.9 to 9.8	(a). From 14.0 to 21.7, then to 14.5	(a). From 26.0 to 18.9, then to 24.5
Outer race fault	(b). From 12.9 to 23.0	(b). From 24.7 to 10.2	(b). From 14.0 to 24.5, then to 19.8	(b). From 25.2 to 14.9, then to 19.5
	(c). From 13.3 to 26.3	(c). From 25.4 to 10.3	(c). From 14.2 to 23.4, then to 17.6	(c). From 25.5 to 15.0, then to 19.6

Table 2. DATA OVERVIEW (DATA SIZE) AND CASE FORMULATION

	Speed increasing	Speed decreasing	Speed first increasing, then decreasing	Speed first decreasing, then increasing
Healthy (Class 1)	100 (Case 1)	100 (Case 2)	100 (Case 3)	100 (Case 4)
Inner race fault (Class 2)	100	100	100	100
Outer race fault (Class 3)	100 (Case 1)	100 (Case 2)	100 (Case 3)	100 (Case 4)

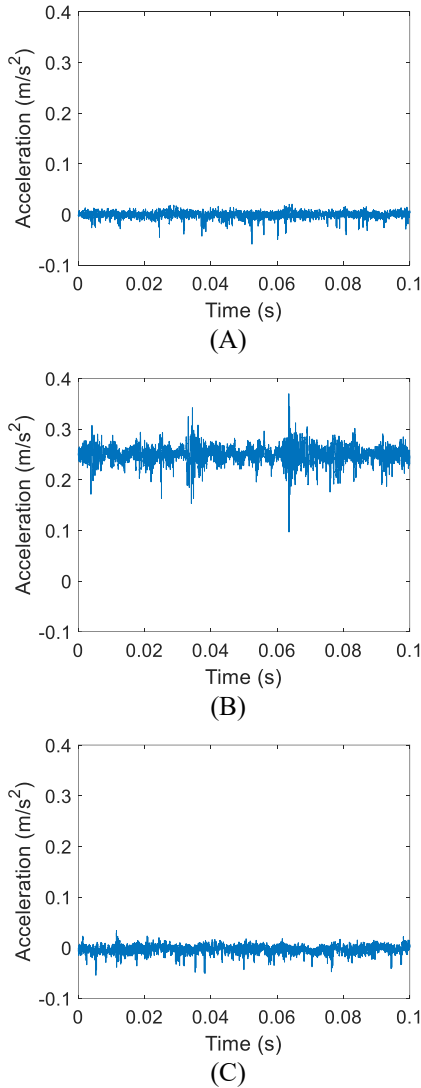


FIGURE 3. TIME-SERIES SAMPLES (A) HEALTHY CONDITION; (B) INNER RACE FAULT; (C) OUTER RACE FAULT.

3.2 DLSTM network establishment and fault diagnosis performance investigation

Recall the capability of LSTM in constructing the temporal correlation of fault with respect to its pivot features over time and the powerful feature extraction of CNN. We establish a DLSTM network by fully harnessing the collective advantages. For validation purpose, we also establish a CNN which is a subpart of DLSTM network (i.e., without the LSTM layer) as can be observed in Figure 4. The fault diagnosis analysis will be implemented using both models and the respective diagnosis accuracy will be investigated and compared. Following the general guideline for deep learning neural network design, the architecture of DLSTM network is configured empirically. This network mainly consists of three convolutional layer stacks (CLS), one LSTM layer and other layers. The layer configuration

details are shown in Table 3. One may notice that the input size is inconsistent with the size of single sample, i.e., 20,000. The reason is that the input layer here takes each time sequence of the sample. In this research, we particularly divide 20,000 data points in the sample into 20 time sequences, and each time sequence thus has 1,000 data points. This requires the input layer to have the size as $1 \times 1,000 \times 1$. It is worth pointing out that, the selection of sequence size plays a role in dictating the diagnosis performance, which can be considered as another hyperparameter subject to tuning. To ensure the stable training, the batch normalization is applied on each convolutional layer. Additionally, ReLU activation function is adopted to map the input-output nonlinearity. A multi-class classification problem is investigated in this research because the output can be each of three fault classes to be identified.

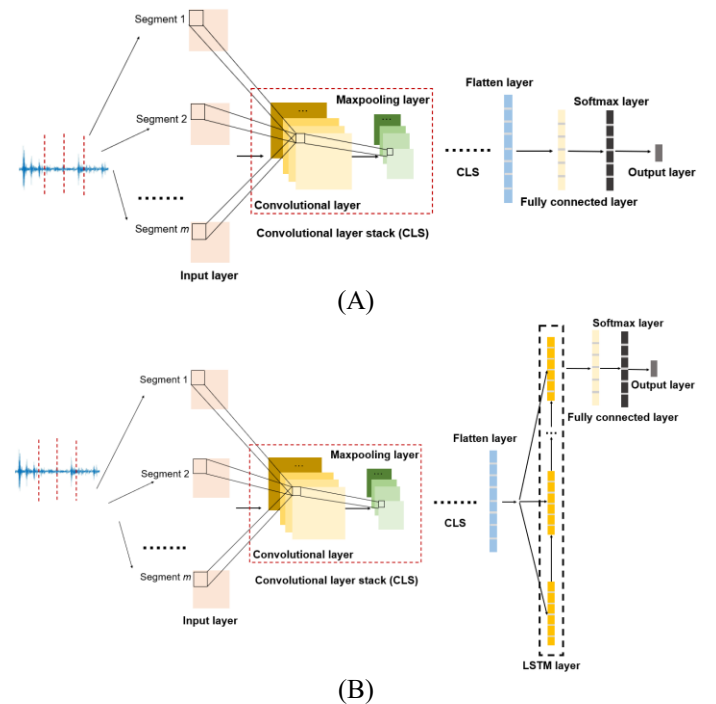


FIGURE 4. MODEL ARCHITECTURE (A) BASELINE CNN MODEL; (B) PROPOSED DLSTM MODEL.

Table 3. LAYER CONFIGURATION

Layer ID	Layer Type	Property	Output Size
1	Sequence Input	N/A	$1 \times 1,000 \times 1$
2	Convolutional	20 filters with size $1 \times 5 \times 1$, same padding	$1 \times 1,000 \times 20$
3	Maxpooling	Max pooling 1×3 , no padding	$1 \times 333 \times 20$

4	Convolutional	20 filters with size $1 \times 5 \times 1$, no padding	$1 \times 329 \times 20$
5	Maxpooling	Max pooling 1×3 , no padding	$1 \times 109 \times 20$
6	Convolutional	20 filters with size $1 \times 5 \times 1$, no padding	$1 \times 105 \times 20$
7	Maxpooling	Max pooling 1×3 , no padding	$1 \times 35 \times 20$
8	Flatten	N/A	700
9	LSTM	hidden units	200
10	Fully connected	3 nodes	3
11	Softmax	3 nodes	3
12	Output/classification	1 node	1

Note: the number of total learnable parameters is 723,623.

The hyperparameter tuning is performed according to the network training and validation performance, resulting in the finalized simulation parameters given in Table 4. Such simulation parameters can avoid both underfitting and overfitting. With the well-configured DLSTM network, the emulations corresponding to previously formulated cases are carried out. Because of the randomness in model training and data split [19,20], we specifically implement 10 emulation runs for each testing case and summarize the statistical results of all testing cases, including the classification accuracy and loss (Figures 5 and 6). Noteworthy, loss here is formulated as the categorical cross-entropy which is a common metric in classification analysis. To be precise, the loss shown in Figure 6 is the mean of losses calculated upon 200 testing samples. Such statistical results allow one to comprehensively examine the overall performance and robustness of the proposed methodology.

Table 4. SIMULATION PARAMETERS

Optimizer	Batch size	Epoch size	Learning rate
Adam	5	10	0.0001

As can be seen, DLSTM notably outperforms CNN in terms of both accuracy and loss. Specifically, nearly all emulations of DLSTM can yield 100% accuracy in all testing cases. In comparison, the inferior accuracy of CNN is observed. The wider distributions of accuracy also indicate the unrobust diagnosis performance. The accuracy improvement of DLSTM especially is more significant in Cases 1 and 4 than that in Cases

2 and 3. This may be because that the temporal dependency of features accounting for the effect of the time-varying conditions in Cases 1 and 4 is much stronger. Loss information provides the consistent observation as accuracy. The statistical results in Figures 5 and 6 clearly illustrate the feasibility of the proposed methodology in identifying the bearing fault under time-varying conditions.

We also examine the emulation runs of CNN that yield the worst accuracy, from which the insight regarding the misclassification can be gained. Two emulation runs pointing to Cases 1 and 4 respectively are identified and the associated confusion matrix information is provided in Figure 7. They both indicate that the outer race fault will be easily misclassified into other fault types when the rotational speed varies over time. The fault-related features of outer race fault appear to resemble that of inner race fault in the case of increasing rotational speed (Case 1) since the misclassification is dominated by the inner race fault. The different observation however is captured in Case 4, in which both healthy condition and inner race fault are likely to be the misclassified fault labels. By incorporating LSTM layer into the network to account for the effect of time-varying conditions, the misclassification can be significantly alleviated.

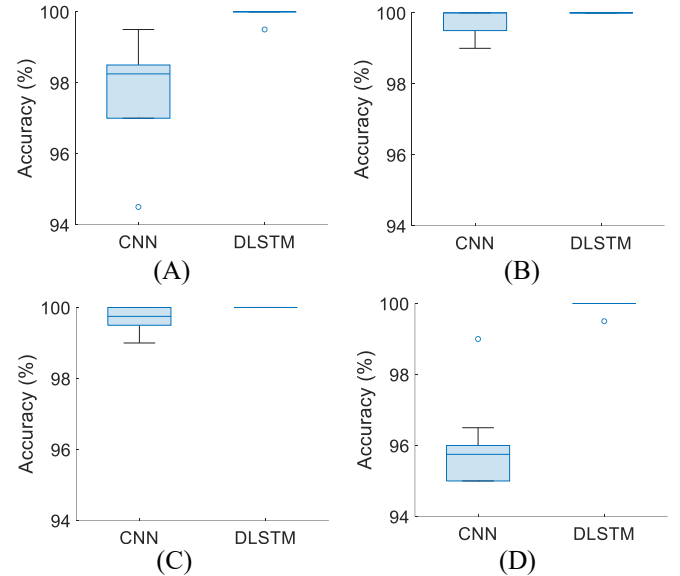
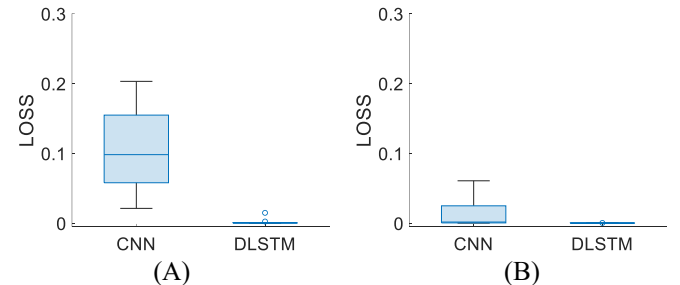


FIGURE 5. ACCURACY COMPARISON (A) CASE 1; (B) CASE 2; (C) CASE 3; (D) CASE 4.



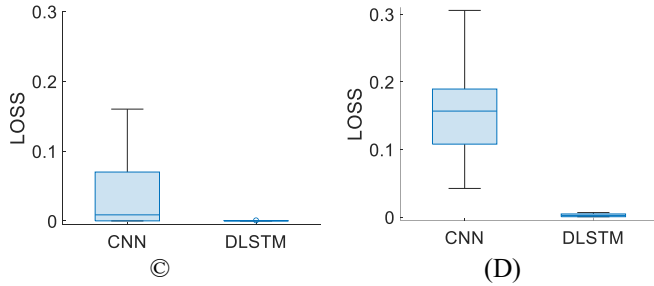


FIGURE 6. LOSS COMPARISON (A) CASE 1; (B) CASE 2; (C) CASE 3; (D) CASE 4.

Output Class	1	100 50.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	0 0.0%	11 5.5%	0.0% 100%
	3	0 0.0%	0 0.0%	89 44.5%	100% 0.0%
		100% 0.0%	NaN% NaN%	89.0% 11.0%	94.5% 5.5%
		Target Class			

(A)

Output Class	1	100 50.0%	0 0.0%	4 2.0%	96.2% 3.8%
	2	0 0.0%	0 0.0%	3 1.5%	0.0% 100%
	3	0 0.0%	0 0.0%	93 46.5%	100% 0.0%
		100% 0.0%	NaN% NaN%	93.0% 7.0%	96.5% 3.5%
		Target Class			

(B)

FIGURE 7. CONFUSION MATRIX INFORMATION (A) WORST EMULATION IN CASE 1; (B) WORST EMULATION IN CASE 4. (CLASS 1: HEALTHY CONDITION; CLASS 2: INNER RACE FAULT; CLASS 3: OUTER RACE FAULT)

3.3 Discussion of future research endeavor

It is evident through the case illustrations that the DLSTM can ensure the desired fault diagnosis performance (i.e., 100% classification accuracy) for this particular dataset because of the relatively small number of fault labels involved in the dataset. Nevertheless, in practical scenarios more fault types will be considered especially when some fault types exhibit the continuous severity. Their associated features hence will become difficult to be discriminated, which will be further compounded by the time-varying conditions. To tackle this issue, the multi-sensor fusion appears to be a potential technique, where the rotational speed measured via encoder or tachometer is further incorporated. There exist two primary means to achieve such sensor fusion. The first one is to design the network architecture with appropriate input layers to feed both acceleration and speed information simultaneously, upon which the proper data fusion strategy will be proposed accordingly. The other one is to resort to the signal processing analysis by fully utilizing both the measured speed and acceleration information. Aiming at reducing the adverse impact of time-varying conditions, one of well-known approaches that is so called the synchronous averaging can be performed either in time domain or time-frequency domain. For illustration, Figure 8 gives the time-frequency synchronous averaging procedures based upon the multi-sensor fusion, which is subject to the future research.

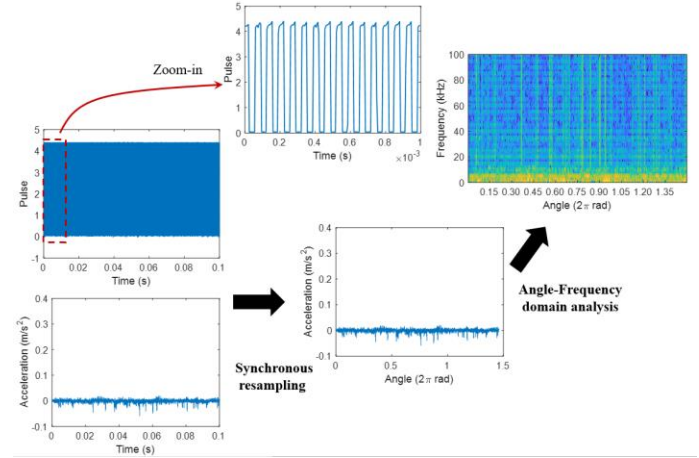


FIGURE 8. SIGNAL PROCESSING BASED ON MULTI-SENSOR FUSION.

4. CONCLUSION

In this research, a deep long short-term memory (DLSTM) network is developed to conduct the bearing fault diagnosis under time-varying operational speed conditions. To illustrate its particular advantage in characterizing the intrinsic correlation between the bearing fault and features over time, a publicly accessible dataset that was acquired under various time-varying conditions is utilized. A counterpart of DLSTM network, i.e., CNN is constructed, and its performance is used as a baseline. Different testing cases are formulated, upon which the statistical validation procedures are executed to facilitate the thorough performance assessment. The results in case studies clearly indicate that the proposed DLSTM network outperforms the CNN in all testing cases in terms of classification accuracy, showing its effectiveness for fault diagnosis under time-varying conditions. Multi-sensor fusion is considered as one future research direction to further enhance the diagnosis performance.

ACKNOWLEDGMENT

This research is supported by National Science Foundation under grant CMMI-2138522.

REFERENCES

- [1] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, F. Fnaiech, Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals, *Appl. Acoust.* 89 (2015) 16–27. <https://doi.org/10.1016/j.apacoust.2014.08.016>.
- [2] L. Song, H. Wang, P. Chen, Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery, *IEEE Trans. Instrum. Meas.* 67 (2018) 1887–1899. <https://doi.org/10.1109/TIM.2018.2806984>.
- [3] M. He, D. He, A new hybrid deep signal processing approach for bearing fault diagnosis using vibration

- signals, *Neurocomputing*. 396 (2020) 542–555. <https://doi.org/10.1016/j.neucom.2018.12.088>.
- [4] Z.K. Peng, P.W. Tse, F.L. Chu, A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing, *Mech. Syst. Signal Process.* 19 (2005) 974–988. <https://doi.org/10.1016/j.ymssp.2004.01.006>.
- [5] J. Chen, J. Pan, Z. Li, Y. Zi, X. Chen, Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals, *Renew. Energy*. 89 (2016) 80–92. <https://doi.org/10.1016/j.renene.2015.12.010>.
- [6] D. Wang, Y. Zhao, C. Yi, K.-L. Tsui, J. Lin, Sparsity guided empirical wavelet transform for fault diagnosis of rolling element bearings, *Mech. Syst. Signal Process.* 101 (2018) 292–308. <https://doi.org/10.1016/j.ymssp.2017.08.038>.
- [7] J. Zheng, H. Pan, J. Cheng, Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines, *Mech. Syst. Signal Process.* 85 (2017) 746–759. <https://doi.org/10.1016/j.ymssp.2016.09.010>.
- [8] Y. Li, Y. Yang, X. Wang, B. Liu, X. Liang, Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine, *J. Sound Vib.* 428 (2018) 72–86. <https://doi.org/10.1016/j.jsv.2018.04.036>.
- [9] M. Liang, K. Zhou, Probabilistic bearing fault diagnosis using Gaussian process with tailored feature extraction, *Int. J. Adv. Manuf. Technol.* (2021). <https://doi.org/10.1007/s00170-021-08392-6>.
- [10] Z. Chen, A. Mauricio, W. Li, K. Gryllias, A deep learning method for bearing fault diagnosis based on Cyclic Spectral Coherence and Convolutional Neural Networks, *Mech. Syst. Signal Process.* 140 (2020) 106683. <https://doi.org/10.1016/j.ymssp.2020.106683>.
- [11] Z. Zhu, G. Peng, Y. Chen, H. Gao, A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis, *Neurocomputing*. 323 (2019) 62–75. <https://doi.org/10.1016/j.neucom.2018.09.050>.
- [12] D.-T. Hoang, H.-J. Kang, Rolling element bearing fault diagnosis using convolutional neural network and vibration image, *Cogn. Syst. Res.* 53 (2019) 42–50. <https://doi.org/10.1016/j.cogsys.2018.03.002>.
- [13] P. Ma, H. Zhang, W. Fan, C. Wang, G. Wen, X. Zhang, A novel bearing fault diagnosis method based on 2D image representation and transfer learning-convolutional neural network, *Meas. Sci. Technol.* 30 (2019) 055402. <https://doi.org/10.1088/1361-6501/ab0793>.
- [14] K. Zhou, J. Tang, Harnessing fuzzy neural network for gear fault diagnosis with limited data labels, *Int. J. Adv. Manuf. Technol.* 115 (2021) 1005–1019. <https://doi.org/10.1007/s00170-021-07253-6>.
- [15] K. Zhou, Y. Liu, Early-Stage Gas identification using convolutional long short-term neural network with sensor array time series data, *Sensors*. 21 (2021). <https://doi.org/10.3390/s21144826>.
- [16] X. Ma, Z. Tao, Y. Wang, H. Yu, Y. Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, *Transp. Res. Part C Emerg. Technol.* 54 (2015) 187–197. <https://doi.org/10.1016/j.trc.2015.03.014>.
- [17] X. Li, X. Wu, Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition, in: 2015 IEEE Int. Conf. Acoust. Speech Signal Process., IEEE, 2015: pp. 4520–4524. <https://doi.org/10.1109/ICASSP.2015.7178826>.
- [18] H. Huang, N. Baddour, Bearing vibration data collected under time-varying rotational speed conditions, *Data Br.* 21 (2018) 1745–1749. <https://doi.org/10.1016/j.dib.2018.11.019>.
- [19] K. Zhou, H. Sun, R. Enos, D. Zhang, J. Tang, Harnessing deep learning for physics-informed prediction of composite strength with microstructural uncertainties, *Comput. Mater. Sci.* 197 (2021) 110663. <https://doi.org/https://doi.org/10.1016/j.commatsci.2021.110663>.
- [20] K. Zhou, J. Tang, Efficient characterization of dynamic response variation using multi-fidelity data fusion through composite neural network, *Eng. Struct.* 232 (2021) 111878. <https://doi.org/https://doi.org/10.1016/j.engstruct.2021.111878>.