Bidirectional Long Short-term Memory Network for Maglev Bridge Acceleration Data Reconstruction

GAO-FENG JIANG, SU-MEI WANG and YI-QING NI

ABSTRACT

Maglev is a developing transportation mode in the recent years, and its safety and stability are required to be ensured by structural health monitoring system. However, the data measured by some sensors may be partially or totally unavailable due to the external electromagnetic interference or sensor failure. While, in some cases, those missing data are important to identify and analyze damage of maglev bridges. Since the structural dynamic characteristics of maglev bridge are complicated, the data reconstruction of those sensors becomes challenging. Therefore, this study proposes a bidirectional long short-term memory (BiLSTM) network to accurately reconstruct the maglev bridge acceleration data. The design of long short-term memory helps to sufficiently exploit the time series data and learn the hidden features of maglev bridge acceleration data between sensors. The bidirectional architecture enables the network to simultaneously learn the time series data from past and future, which is beneficial to extract more hidden features. A dataset collected from an in-site experiment for maglev bridge is used to verify the feasibility of the proposed method. The lost acceleration data from abnormal sensors is predicted by the acceleration data recorded from normally operating sensors. The results shows that the difference between predicted and true acceleration data is at a very low magnitude. Consequently, the proposed method can be applied for the high-performance reconstruction of maglev bridge acceleration data.

INTRODUCTION

Maglev system consists of train, suspension bogie, rail, and bridge. The degradation of maglev bridges may affect the structural stability of maglev system. To control the structural performance of maglev systems and improve the reliability of bridges, it is critical to predict the acceleration from different sensors installed on the maglev bridge by establishing a fair and accurate predictive model, which is the basis for maglev bridge acceleration predictions.

With the thriving progress in artificial intelligence and big data technology, deep learning methods have been widely applied in the modeling of maglev bridge acceleration prediction. Since the maglev bridge accelerations are a kind of time-series, recurrent neural networks (RNNs), as a class of the deep learning methods, are promising to realize the data-driven predictive model in the maglev bridge acceleration. In RNN, the transmission of information is shaped as a loop by allowing the output from one unit to influence the subsequent input to the same unit. Hence, this so-called feedback connections in RNN enable the prediction on an entire time-series.

However, the long-term gradients which are backpropagated can vanish or explode thus RNN may be unable to transmit the information efficiently. To solve this shortage, long short-term memory (LSTM) network is introduced [1]. As a special design of RNN, LSTM consists of three gates responsible for the transmission of information. Due to this, LSTM is outperformed than vanilla RNNs on some temporal data processing tasks [2].

Gao-Feng Jiang, Su-Mei Wang, Yi-Qing Ni, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong S.A.R., China.

As a kind of special LSTM, this study aims to propose a bidirectional LSTM (BiLSTM) network for the prediction of multiple maglev bridge accelerations. First, some acceleration sensors collect historical acceleration data on the maglev bridges, followed by the acquisition of dataset prepared for BiLSTM network. Then, a portion of them is manipulated as the lost data of maglev bridge accelerations, and the remaining acceleration data is regarded as the training dataset. Finally, the prediction is realized by BiLSTM network for the data reconstruction of lost accelerations for a maglev bridge.

METHODOLOGY

Since LSTM is a special type of RNN, the definition of RNN is first given. Fig. 1 shows the structure of RNN, where unit A is the main body that receives the input x_t and generates the output h_t. The loop structure of RNN enables the information of one moment to be transmitted to the next moment. Thus, the information can be retained by RNN, which is impossible in other neural networks. As shown in Fig. 1, RNN can be thought of being made up of multiple copies of the identical unit, and the output of each unit serves as the input of its following unit. The structure of RNN is like a chain through the time, which can be regarded as a natural design for a neural network to exploit time series data.

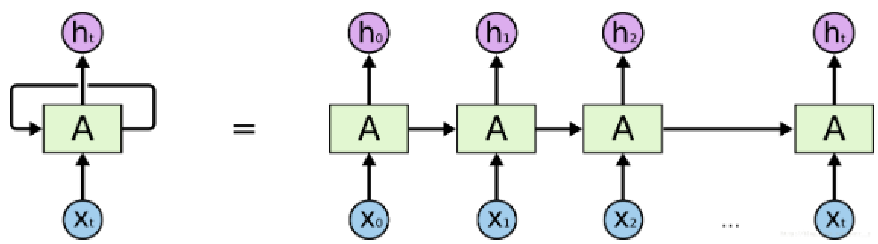
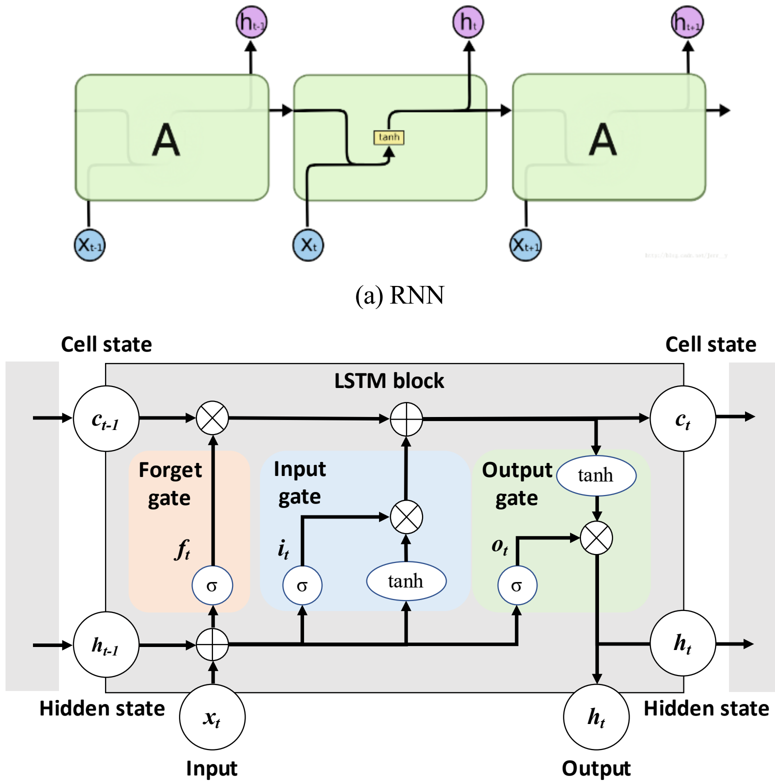


Figure 1. The structure of recurrent neural network.

Fig. 2 shows the design of block structure between unit of LSTM and RNN. Both have the cell for remembering the values over arbitrary time intervals, this helps both to form a short-term memory for the information. However, the unit consists of only a very simple layer (such as a single tanh function) in RNNs while multiple layers in LSTM. The layer difference lies in the design of sigmoid function σ, which is a lack in RNN. The output of each sigmoid function represents the weight on the transmission of information and varies from 0 to 1, which means that the transmission of information is totally prohibited when $\sigma = 0$ while is totally permitted when $\sigma = 1$.

The sigmoid function is also called gate in LSTM. From the left to right in a LSTM unit, the name of three gates is forget gate, input gate and output gate, respectively. Assumed the input of a LSTM unit is denoted as x_t, the former state and current state of cell is denoted as c_{t-1} and c_t, and the output of a LSTM is denoted as h_t, the goal of three gates are that

Forget gate: determination on the amount of information dropping from h_{t-1} and c_{t-1},
Input gate: determination on the amount of information adding from x_t to c_t,
Output gate: determination on the amount of information updating h_t.



(a) RNN

(b) LSTM

Figure 2. The block structure in RNN and LSTM.

To realize the function of forget gate, input gate, and output gate, the following equations are satisfied:

$$i_t = \sigma(W_i x_t + W_{hi} h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_f x_t + W_{hf} h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(W_o x_t + W_{ho} h_{t-1} + b_o) \quad (3)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh\left(W_g x_t + W_{hg} h_{t-1} + b_g\right) \quad (4)$$

$$h_t = o_t \odot \tanh(c_t) \quad (5)$$

By using these three gates, LSTM can realize the control on the transmission of information in both long-term and short-term memory. Thus, this is well-suited as the predictive model for time-series. LSTM has been widely utilized by the researchers for purposes of structural identification and response prediction [3-5].

The regular LSTM network allows the information flow in only one direction. While in the bidirectional LSTM (BiLSTM) network, the information can be transmitted in two directions, i.e., in backwards and forwards, as represented in Fig. 3. The hidden states computed from two information flows are concatenated. As a result, the information from both past and future states can be obtained simultaneously, which effectively increase the amount of information available. Applications of BiLSTM in structural engineering can be found in the recent literature [6-8] and deserve more studies.

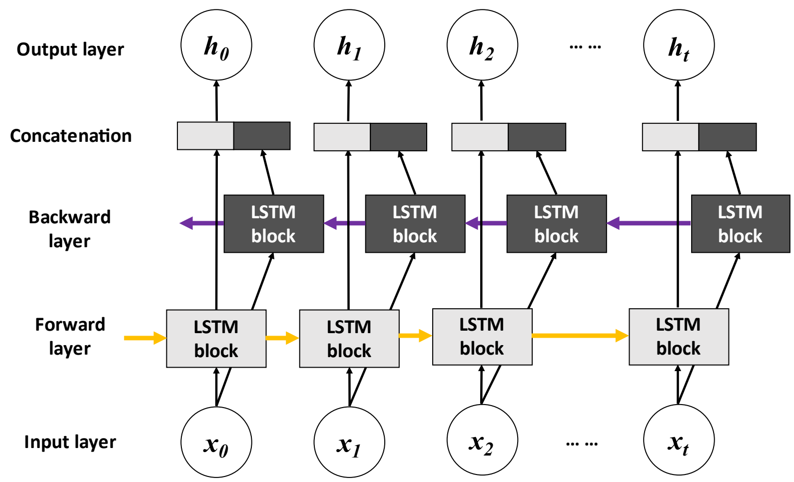


Figure 3. The architecture of bidirectional LSTM (BiLSTM) network.

RESULTS

Data preparation

The dataset is originated from an experimental test conducted on Shanghai Medium Speed Maglev Line. Several accelerometers are installed to detect the acceleration of maglev bridges, and all the accelerometers are used to collect vertical accelerations. The data is measured by a 16-channel data acquisition equipment and recorded in a high-memory and high-storage portable computer. The sampling rate of data acquisition is equal to 5,000 Hz. Four accelerometers, denoted A1 to A4, are representative of this report. A1 and A2 is installed on the steel bridge, where A1 is at quarter location of bridge and A2 is at middle location of bridge. A3 and A4 is installed on the concrete bridge, where A3 is at quarter location of bridge and A4 is at middle location of bridge. The dataset is packaged from five individual groups, and each group simultaneously covers the acceleration data from A1 to A4. The length of acceleration data equals to 100,000, meaning that each group records the change of acceleration in a total of 20 seconds.

Model establishment

The proposed BiLSTM model contains an input layer, a LSTM layer, and an output layer. The input size and output size are both equal to 1, thus this study is a single-to-single data reconstruction. The adjustable hyperparameters are defined that the size of the hidden state is 40, and the learning rate of model is 0.01. The model is optimized by using Mean Square Error (MSE), which is also the criterion of model assessment. The model is trained under 1,000 iterations, and the programming is implemented with PyTorch 1.12.1 and Python 3.8.

Data reconstruction

The acceleration data is reconstructed from A1 to A2, A2 to A1, A3 to A4 and A4 to A3. Therefore, four tasks are achieved in this study, denoted task A to task D. Data from each group is used to establish a specified BiLSTM model in each task. Only first 0.1 second data (data length is 500) is selected as training dataset to preserve the high model prediction performance with a relatively low time cost of model training.

Table I shows the results of model performance, which are represented as MSEs between ground truth and prediction obtained from five groups. The MSEs are calculated in a range of total length, thus it reflects the predicting generalizability of BiLSTM model. Overall, the MSEs are at a very low magnitude of 10^{-5}, though the training data length is only hundredth of whole data length. The MSEs from task C and D are much lower than that from task A and B. It means that the acceleration data from concrete bridge is easier to be extracted its features. It is close of MSEs between task C and task D, while MSEs from task A are much larger than that from task B. It indicates that the data reconstruction is more difficult for steel bridge compared to concrete bridge. There exists a difference of MSEs among five groups. The prediction performs the best in Group V, where the MSEs in all four tasks are the lowest in this group. Task C and task D performs the worst in Group IV, while task A and task B performs the worst in Group III.

TABLE I. MODEL PERFORMANCE IN FOUR TASKS (UNIT: $\times 10^{-5}$)

	Group I	**Group II**	**Group III**	**Group IV**	**Group V**
Task A	40.531	18.702	70.320	10.846	8.850
Task B	4.214	9.402	25.892	4.054	2.526
Task C	0.595	1.061	1.044	6.640	0.318
Task D	0.479	0.999	0.865	7.612	0.275

To straightforwardly demonstrate the prediction effectiveness by using BiLSTM model, Fig. 4 shows the ground truth and prediction results extracted from Group V.

As shown in Fig. 4, the good agreement between ground truth and prediction is found from task C and task D, while there is a huge deviation between ground truth and prediction in task A and task B. This results in a line with the MSE results obtained from Table I. To make a better prediction, the proposed BiLSTM model should be adjusted with its parameters to satisfy the requirements of data reconstruction in task A and task B.

To illustrate the superiority of using BiLSTM, Table II shows the MSEs of group V obtained from three methods, that are BiLSTM, LSTM and RNN. Task C and task D are adopted since the features from time-series can only learnt from them. In both tasks, BiLSTM owns the best performance with the lowest MSEs, followed by LSTM and RNN. Especially in task D, the existence of LSTM block boosts the model performance when comparing RNN to LSTM and BiLSTM. Therefore, BiLSTM is expected to be the best option for data reconstruction in this study.

TABLE II. PERFROMANCE OF THREE METHODS IN TASK C AND TASK D (UNIT: $\times 10^{-5}$)

	RNN	LSTM	BiLSTM
Task C	0.328	0.319	0.318
Task D	0.400	0.282	0.275

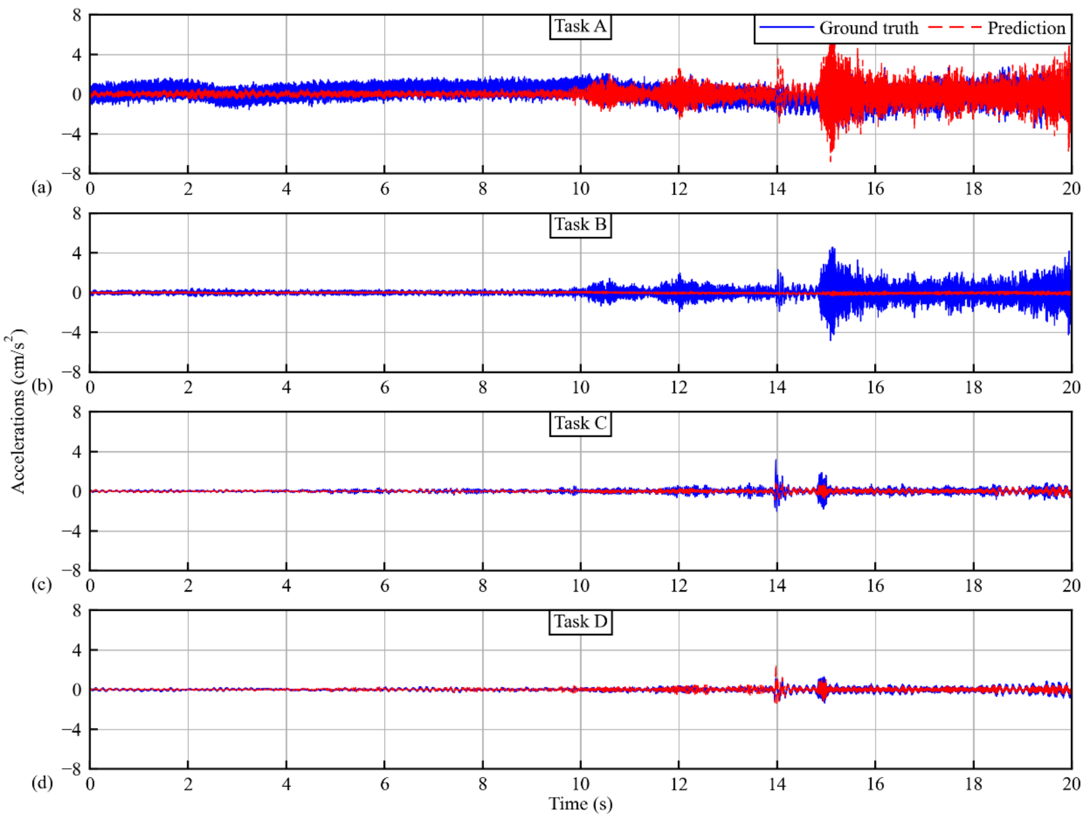


Figure 4. The ground truth and prediction of whole-length data in Group V.

CONCLUSION

The structural health monitoring relies on a comprehensive dataset to make a data-driven analysis for condition monitoring, damage detection, and risk prediction. However, it is always a great concern that the data is in absence or in a lack of information. To supplement the shortage of data missing, this study proposes a bidirectional long short-term (BiLSTM) memory network to make up those missing data. The effectiveness of method is verified by using acceleration data from bridges under a maglev test line. Results show that the use of BiLSTM can realize the objective of data reconstruction and may receive a better prediction compared to LSTM and RNN. It is available for reconstructing data from a concrete bridge, and there is still a room of improvement for reconstructing data from a steel bridge. Further study will focus on the adaptability of method, the increase of prediction accuracy, and the extension of application scenarios.

REFERENCES

1. Hochreiter, S., and Schmidhuber, J. 1997. "Long short-term memory," *Neural Comput.*, 9(8): 1735-1780.
2. Gers, F. A., Schmidhuber, J., and Cummins, F. 2000. "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, 12(10): 2451-2471.
3. Wang, Y. 2017. "A new concept using LSTM neural networks for dynamic system identification, in *2017 American Control Conference (ACC)*.
4. Zhang, R., Chen, Z., Chen, S., Zheng, J., Büyüköztürk, O., and Sun, H. 2019, "Deep long short-term memory networks for nonlinear structural seismic response prediction", *Comput. Struct.*, 220: 55–68.
5. Yue, Z., Ding, Y., Zhao, H., and Wang, Z. 2022. "Mechanics-guided optimization of an LSTM network for real-time modeling of temperature-induced deflection of a cable-stayed bridge", *Eng. Struct.*, 252: 113619.
6. Yoon, D., Yech, Z., and Byun, J. 2020. "Seismic data reconstruction using deep bidirectional long short-term memory with skip connections", *IEEE Geosci. Remote Sens. Lett.*, 18(7): 1298–1302.
7. Tian, Y., Xu, Y., Zhang, D., and Li, H. 2021. "Relationship modeling between vehicle-induced girder vertical deflection and cable tension by BiLSTM using field monitoring data of a cable-stayed bridge", *Struct. Control Health Monit.* 28(2): e2667.
8. Lu, Y., Tang, L., Chen, C., Zhou, L., Liu, Z., Liu Y., and Yang, B. 2023. "Reconstruction of structural long-term acceleration response based on BiLSTM networks. *Eng. Struct.* 285:116000.