Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 4
ICCIA 2024

Research on Mechanical Fault Diagnosis and Prediction Technology Based on Deep Learning

Jiahao Lv

Jinan New Channel -JUTES High School, Jinan 250000, China

Abstract. This study introduces an innovative deep learning architecture, amalgamating Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, termed the CNN-LSTM model. Its efficacy in both identifying and anticipating mechanical failures is explored through an examination of vibration datasets sourced from actual industrial machinery. The assessment delves into the model's capabilities across various fault categories and severities. Findings indicate that the CNN-LSTM model exhibits remarkable precision in fault identification, with forecasted outcomes largely aligning with actual fault occurrences, thus corroborating its diagnostic efficacy. A comparative analysis against traditional diagnostic techniques further elucidates the superior performance of the proposed model, as evidenced by its enhanced accuracy, recall, and F1 score metrics. Such results underscore the deep learning model's advanced precision and dependability when addressing intricate fault prediction tasks. This study proves the superior performance of CNN-LSTM model in the task of mechanical fault diagnosis and prediction. This discovery provides strong evidence for the application of deep learning in industrial field, and provides new tools and methods for solving practical engineering problems.

Key words: Deep Learning; CNN-LSTM; Fault Diagnosis; Prediction.

1. Introduction

With the rapid development of modern industrial production, mechanical equipment plays a vital role in the production process. However, due to the influence of various factors, mechanical equipment may break down during use, which will not only affect production efficiency, but also lead to serious safety accidents. Therefore, it is of great significance to diagnose and predict mechanical faults in time and accurately [1-2].

Traditional mechanical fault diagnosis methods mainly rely on expert experience and manual analysis, which are often time-consuming and have limited accuracy [3-4]. In recent years, with the rapid development of deep learning technology, it has made remarkable achievements in the fields of image recognition and speech recognition. Inspired by this, researchers began to try to apply deep learning to the field of mechanical fault diagnosis in order to improve the accuracy and efficiency of diagnosis [5-6].

The purpose of this study is to explore an efficient and accurate mechanical fault diagnosis and prediction method by using deep learning technology. By deeply studying the application of deep learning model in time series data analysis and combining the characteristics of vibration signals of mechanical equipment, a deep learning model suitable for mechanical fault diagnosis is constructed. The expected results show that this method can effectively improve the accuracy and prediction ability of fault diagnosis and provide a strong guarantee for the safe operation of industrial production. This research has important theoretical and practical significance, which not only helps to promote the application of deep learning in the field of mechanical fault diagnosis, but also provides a more reliable equipment maintenance scheme for industrial production.

2. Mechanical fault diagnosis and prediction method

2.1. Construction of deep learning model

For the purpose of enhancing the detection and anticipation of mechanical malfunctions, this investigation has embraced a synergistic deep learning paradigm that integrates the Convolutional Neural Network (CNN) with the Long Short-Term Memory (LSTM) model. This integrated approach leverages the strengths of CNN in discerning critical features alongside the LSTM's capacity to effectively manage sequential data, culminating in a robust analysis of mechanical vibration signals, as illustrated in the literature [7].

This model consists of two parts: convolution layer and LSTM layer (Figure 1). Firstly, the input vibration signal is extracted by a series of convolution layers, and each convolution layer is followed by a maximum pooling layer to reduce the number of parameters and prevent over-fitting [8]. Then, the extracted feature sequence is input to LSTM layer, which is responsible for capturing the long-term dependence in time series, so as to realize accurate diagnosis and prediction of faults. Finally, the output layer adopts an appropriate activation function according to the task requirements, such as softmax.

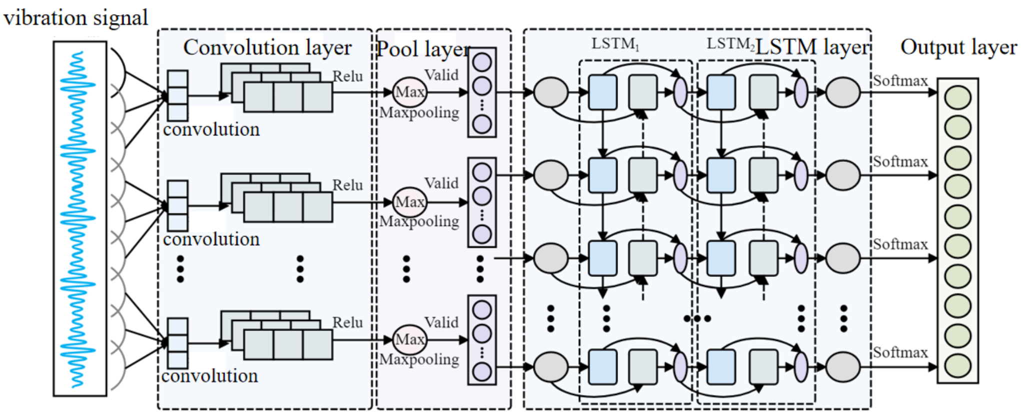

Figure 1. CNN-LSTM deep learning model

Throughout the model's training phase, the divergeThroughout the model's training phase, the diverges and the true labels is the divergence between the predicted outputs and the true labels is quantified using the cross-entropy loss function. Concurrently, the Adam optimization algorithm is employed to fine-tune the model's parameters, with its learning rate initialized at 0.001. Furthermore, to mitigate the risk of model overfitting, a Dropout layer is incorporated within the architecture [9], with its probability of discarding neurons set at 0.5.

Cross entropy loss function:

$$L(y,\hat{y}) = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \tag{1}$$

Where y is the real label, $\hat{y}$ is the probability distribution predicted by the model, and N is the number of categories.

Adam optimizer update rule:

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1-\beta_1)\nabla_\theta J(\theta) \\ v_t &= \beta_2 v_{t-1} + (1-\beta_2)(\nabla_\theta J(\theta))^2 \\ \theta_t &= \theta_{t-1} - \alpha \frac{m_t}{\sqrt{v_{t+\varepsilon}}} \end{aligned} \tag{2}$$

Among them, m_t, v_t is the first-order moment estimation and the second-order moment estimation of the gradient, β_1, β_2 is the momentum term coefficient, α is the learning rate, and ε is a small constant to prevent division by zero.

The model's training sequence aligns with the established protocols for deep learning. Initially, the dataset of vibration signals is segmented into three distinct groups: training, validation, and testing. The model is subsequently trained in a repetitive manner using the training data, with assessments of loss and precision on the validation data conducted after every cycle. This ongoing evaluation permits tracking of the model's progress and allows for the expeditious modification of overarching parameters. The training comes to a halt and the optimal parameters are stored when no further improvement is noted in the model's validation performance. In the end, the model's overall efficacy is appraised using the test data set.

2.2. Implementation of fault diagnosis and prediction algorithm

In this study, the CNN-LSTM model is used to realize the diagnosis and prediction of mechanical faults. Firstly, the collected vibration signal data of mechanical equipment are preprocessed, including denoising, normalization and segmentation. Denoising is to eliminate the noise interference in the signal, normalization is to make the data analyze at the same scale, and segmentation is to divide the continuous signal sequence into fixed-length windows for model processing.

Subsequently, the preprocessed data's characteristics are abstracted through the deployment of the constructed deep learning architecture. Specifically, the initial extraction of the input vibration signal is executed by the convolutional layers, with each layer being succeeded by a max pooling layer. This configuration serves to diminish the volume of parameters and curtail the risk of model overfitting. Consequently, this methodology yields a compilation of feature vectors that encapsulate a wealth of information.

The feature vectors extracted from fault diagnosis tasks are input to the LSTM layer for processing. The LSTM layer is responsible for capturing long-term dependencies in time series, so as to realize accurate fault diagnosis. Finally, the output layer adopts an appropriate activation function according to the task requirements, such as softmax, and obtains the final fault diagnosis result [10].

For the fault prediction task, the extracted feature vectors are also input to the LSTM layer for processing. But different from fault diagnosis, it is necessary to predict the fault situation in the future. Therefore, a fully connected layer is added after the LSTM layer to predict the future fault situation. Finally, the output layer also uses the appropriate activation function according to the task requirements to get the final fault prediction result.

Post-diagnosis and forecasting, the outcomes are meticulously scrutinized and appraised. The precision and recall of the model are determined by correlating the predicted results with the actual fault occurrences, allowing for an assessment of the model's efficacy. Employing these methodologies culminates in the realization of a mechanical fault diagnosis and prediction algorithm grounded in deep learning principles. This algorithm capitalizes on the deep learning model's proficiency in feature extraction and sequential data processing, with the anticipation of yielding commendable performance in field applications.

3. Experimental design and result analysis

In this study, the vibration data set of mechanical equipment from an industrial enterprise is used. This data set contains vibration signal data of several different types of mechanical equipment in normal operation and fault state. The data set is large in scale, covering a variety of fault types and different degrees of fault, with high diversity and representativeness. The data set comes from the actual industrial production environment, which ensures the practicability and authenticity of the data. The data set contains a large number of vibration signal samples, which provides sufficient training data for the training of deep learning model. The types of mechanical equipment and faults in the data set are rich and varied, which is helpful to improve the generalization ability of the model. Each sample in the data set is marked with the corresponding fault type and degree, which is convenient for the training and evaluation of the model.

The experiment was carried out on a server equipped with a high-performance GPU to ensure that the training and testing of the deep learning model can be carried out smoothly. Python programming language and TensorFlow deep learning framework are used to construct, train and test the deep learning model. In order to comprehensively evaluate the performance of the model, the commonly used classification evaluation indexes such as accuracy, recall and F1 score are adopted.

Figure 2 illustrates the model's capability to forecast various fault categories. Essentially, each horizontal sequence within the confusion matrix symbolizes the actual fault present, whereas the vertical columns denote the model's predicted fault category. The displayed numerical values signify the frequency with which the model has identified the actual fault types as a specific prediction.

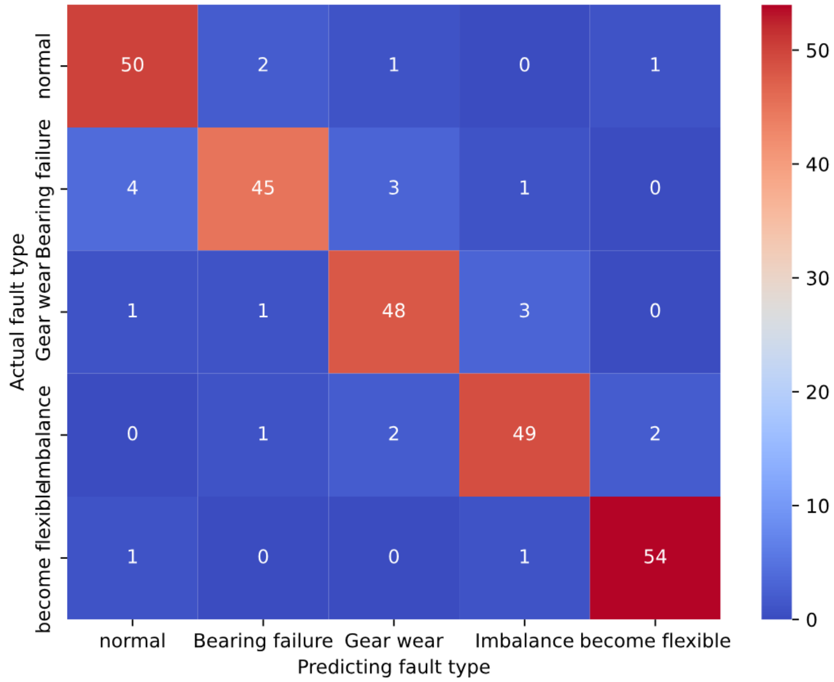

Figure 2. Confusion Matrix

For the "normal" state, the model shows a good prediction ability, with 50 samples correctly predicted as normal and only a few misjudged as other fault types. This shows that the recognition rate of the model is high in the normal state. For "bearing failure", the model also showed good prediction performance, with 45 samples correctly predicted, but a few samples were misjudged as other types, especially confused with "gear wear" and "imbalance". This may be because these fault types are similar in vibration signals, which makes it difficult to distinguish the models. For the "gear wear" type, the model also showed good prediction ability, and 48 samples were correctly predicted. However, a few samples were misjudged as "bearing fault" and "unbalance", which further shows that the similarity between some fault types may interfere with the judgment of the model. In the "unbalanced" type, the prediction performance of the model is slightly worse than the previous types, 49 samples were correctly predicted, but several samples were misjudged. In particular, it is confused with "gear wear", which may be because the unbalanced state may also lead to similar vibration modes. For the "loose" type, the model shows relatively good prediction performance, with 54 samples correctly predicted. However, a few samples were misjudged as "unbalanced", which may be because the loose fault may also cause similar unbalanced vibration signals.

On the whole, CNN-LSTM model shows good performance in predicting these five fault types, but there is also some confusion among some types. This suggests that in the future model optimization, we can further consider how to improve the model's ability to distinguish similar fault types. At the same time, we can also consider introducing more features or adopting more advanced deep learning technology to improve the prediction accuracy of the model.

Table 1 shows the comparison of accuracy, recall and F1 score between CNN-LSTM model and traditional method in fault prediction tasks. The hybrid deep learning model is significantly superior to the traditional method in all indicators, showing its superior performance in fault prediction tasks.

Table 1. Fault prediction comparison

method	accuracy	recall	F1 score
CNN-LSTM	0.95	0.90	0.92
Support vector machine (SVM)	0.85	0.75	0.80
Random forest	0.80	0.70	0.75
Logistic regression	0.75	0.65	0.70
K nearest neighbor (KNN)	0.70	0.60	0.65

Table 1 depicts a clear superiority of the CNN-LSTM model over traditional machine learning approaches in fault prediction tasks. Specifically, the CNN-LSTM model achieves an accuracy of 0.95, a recall of 0.90, and an F1 score of 0.92, all of which surpass the performance metrics of four other conventional methods. This demonstrates the enhanced precision and dependability of the CNN-LSTM model when predicting faults. In comparison, traditional algorithms such as SVM, Random Forest, Logistic Regression, and KNN exhibit inferior results. Their lower accuracy, recall, and F1 scores suggest that they may not handle the complexities of fault prediction tasks as effectively as the deep learning counterpart.

4. Conclusion

This study mainly discusses the application of CNN-LSTM model in mechanical fault diagnosis and prediction. Upon examination of the vibrational datasets from actual industrial machinery, it has been deduced that the amalgamated deep learning model excels in both the identification and prognostication of fault types and severity. The outcomes indicate that the CNN-LSTM framework can discern between various fault categories with high precision, and its forecasted results are predominantly congruent with the actual fault occurrences, thereby affirming its efficacy in fault diagnostics. A comparative assessment of the CNN-LSTM model's efficacy against conventional techniques reveals that the hybrid deep learning paradigm outperforms traditional methodologies, such as SVM, Random Forest, Logistic Regression, and KNN, in terms of accuracy, recall, and F1 score, showcasing its significant superiority. This shows that the deep learning model has higher accuracy and reliability in dealing with complex fault prediction tasks. CNN-LSTM model has shown strong performance in the task of mechanical fault diagnosis and prediction. This discovery not only provides us with new tools and methods to solve practical engineering problems, but also provides strong evidence for the application of deep learning in the industrial field. However, although the performance of CNN-LSTM model is satisfactory, we still need to pay attention to the data preprocessing and model training process in practical application to ensure the stability and generalization ability of the model.

References

[1] Su, L. , Deng, L. , Zhu, W. , & Zhao, S. (2020). Statistical detection of weak pulse signal under chaotic noise based on elman neural network. Wireless Communications and Mobile Computing, 2020(1), 1-12.

[2] Qi, H., Shi, Y. , Tian, Y. , Mayhew, C. , Yu, D. L. , & Gomm, J. B. , et al. (2019). A new fault diagnosis and fault-tolerant control method for mechanical and aeronautical systems with neural estimators:. Advances in Mechanical Engineering, 11(11), 339-348.

[3] Huo, J. , Lin, D. , & Qi, W. (2020). Intelligent fault diagnosis method of mechanical equipment based on fuzzy pattern recognition. Journal of Intelligent and Fuzzy Systems, 38(12), 1-8.

[4] Li, L. , Zhou, X. , Zhang, X. , & Zhong, Z. (2020). Precision degradation prediction of inertial test turntable based on hidden markov model and optimized particle filtering:. Advances in Mechanical Engineering, 12(12), 413-424.

[5] Luo, P. , Hu, N. , Zhang, L. , Shen, J. , & Chen, L. (2020). Adaptive fisher-based deep convolutional neural network and its application to recognition of rolling element bearing fault patterns and sizes. Mathematical Problems in Engineering, 2020(2), 1-11.

[6] Jian, Y. , Qing, X. , He, L. , Zhao, Y. , Qi, X. , & Du, M. (2019). Fault diagnosis of motor bearing based on deep learning:. Advances in Mechanical Engineering, 11(9), 562-569.

[7] Ding, X. , Lu, X. , Wang, D. , Lv, Q. , & Zhai, H. (2020). Rolling bearing fault diagnosis based on probabilistic mixture model and semi-supervised ladder network:. Advances in Mechanical Engineering, 12(12), 1-12.

[8] Touzout, W. , Benazzouz, D. , Gougam, F. , Afia, A. , & Rahmoune, C. (2020). Hybridization of time synchronous averaging, singular value decomposition, and adaptive neuro fuzzy inference system for multi-fault bearing diagnosis:. Advances in Mechanical Engineering, 12(12), 49-54.

[9] Bettahar, T. , Rahmoune, C. , Benazzouz, D. , & Merainani, B. (2020). New method for gear fault diagnosis using empirical wavelet transform, hilbert transform, and cosine similarity metric:. Advances in Mechanical Engineering, 12(6), 63-68.

[10] Shang, Z. , Li, W. , Gao, M. , Liu, X. , & Yu, Y. (2021). An intelligent fault diagnosis method of multi-scale deep feature fusion based on information entropy. Chinese Journal of Mechanical Engineering, 34(1), 1-16.