Fault Diagnosis of Rolling Bearing Based on Vibration Signal and Deep Learning Algorithm

Yimuran AIMAIER[1*]

[1]School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China
*a130980@corro.umm.edu.mx

Abstract. In today's rapidly developing industrial sector, mechanical equipment is also more and more biased towards automation, intelligence. In the process of industrial production, mechanical equipment will inevitably have some failure problems. In view of the traditional manual diagnosis of faults may be excessive consumption of human resources, this paper takes the rolling bearing faults, which are more common in mechanical equipment faults, as the object of research, and proposes a bearing fault diagnosis model based on convolutional neural network by studying the application of deep learning on the vibration signals. This integrated approach enhances diagnostic accuracy by combining vibration signal feature extraction with fault type classification. and enhances the accuracy of fault recognition by training the model. To strengthen the architecture's generalization performance, this paper also proposes to add the method of long and short-term memory network, and study the fault diagnosis model of the dual path of convolutional neural network and long and short-term memory network, which enhances the generalizability, further improves the accuracy of fault recognition, and has a certain degree of applicability, which is convenient for the diagnosis and repair of faults in industrial production.

Keywords: Rolling Bearing, Deep Learning, Neural Network, Fault Diagnosis.

1 Introduction

1.1 Research background

In today's era of rapid industrial development, mechanical equipment has become more and more intelligent and plays a critical role across diverse sectors including petrochemical processing, aerospace engineering, and energy production systems. The growing sophistication of contemporary industrial machinery, its potential failure types and failure mechanisms are more and more difficult to analyze and diagnose from simple empirical formulas or traditional methods, which increases the difficulty of early warning and diagnosis of equipment failure. Failure of critical equipment may cause a series of serious chain problems affecting the whole production process [1]. This requires reliable fault diagnosis techniques to ensure the safe operation of the equipment.

J. Iqbal (ed.), *Proceedings of the 2025 2nd International Conference on Mechanics, Electronics Engineering and Automation (ICMEEA 2025)*, Advances in Engineering Research 272,
https://doi.org/10.2991/978-94-6463-821-9_37

Diverging from classical pattern analysis paradigms, end-to-end learning frameworks eliminate manual feature crafting by discovering intrinsic data regularities through hierarchical abstraction layers, effectively decoding nonlinear correlations within complex signal spaces. It relies on big data to improve the accuracy of model recognition and uses hierarchical abstraction of data features to enhance the model's capability to generalize to unseen examples. Owing to its superior feature extraction capabilities in processing industrial multimodal signals, which fits well with the high-dimensional and nonlinear characteristics of mechanical equipment monitoring data [2], it is of great research significance to explore the deep learning-based mechanical equipment fault warning and diagnosis methods.

The traditional method may sometimes miss some important information of fault detection using signal analysis and processing techniques that include time domain analysis, frequency domain analysis, and time-frequency domain analysis leading to diagnostic errors. On the contrary, modern AI-driven condition monitoring systems employ deep neural architectures to autonomously distil discriminative fault signatures from raw vibration spectra, effectively bypassing manual feature engineering constraints. This method makes full use of the powerful ability of deep learning in feature extraction and pattern recognition, which is especially suitable for processing and analyzing large-scale and complex industrial data, and also can provide higher accuracy with stronger generalization ability. In summary, the deep learning method in has a very broad application prospect. rolling bearing health monitoring and fault diagnosis.

In this research, after understanding the current status of the research on fault diagnosis of mechanical equipment at home and abroad, the rolling bearing is taken as the main object of this topic, and the deep learning method is used to significantly enhance the fault characterization accuracy, which fills the gap in this area.

1.2 State-of-the-art

In recent years, deep learning-based bearing faults classification method is hot application. The Case Western Reserve University's rolling bearing database program was formally established, for rolling system bearing fault diagnosis establishes an innovative analytical framework. The advantage of convolutional neural network system is that it can be quickly mined from a large amount of data to its features, relying on its powerful and reliable feature data extraction and identification of fault classification and prediction capabilities, can be anytime, anywhere more quickly and accurately make identification and prediction of various types of rolling bearing fault types, greatly saving manpower and material resources, improve the safety level of the whole machine production and manufacturing of the level of stability. So far, researchers in international research communities primarily focuses on the most widely used deep learning model structure, such as convolutional neural network, recurrent neural network (RNN), autoencoder (AE), deep belief network (DBN) and so on, a large number of mechanical fault diagnosis methods have been researched and applied fault diagnosis method neural network (CNN) [3]. Zhuyun Chen et al. proposed a two-dimensional based on cyclic spectral coherence (CSCoh) and convolutional). Neural Network (CNN) two-dimensional graph representation based DL fault diagnosis method, which estimates the two-dimensional CSCoh graph of vibration signals through cyclic spectral analysis to provide an orientation discrimination model for specific types of faults, and at the same time adopts the normalization (GN) to normalize the feature graph of the network, which can reduce the internal covariance offset caused by the difference in data distribution [4]. Wang et al. developed a multi-modal fault diagnosis framework (MLMF-CR) for industrial bearing systems, leveraging bidirectional LSTM networks to establish localized diagnostic submodels. Their methodology implements adaptive feature extraction through a preprocessing pipeline involving signal filtration and noise reduction, enabling effective separation of compound fault signatures from concurrent vibration-current sensor streams.

Complementing this approach, Li's team introduced a heterogeneous network architecture integrating CNN and DNN components. The CNN branch processes raw vibration waveforms through learnable filter banks, while the DNN pathway analyzes nine time-domain statistical descriptors (including kurtosis and RMS values), achieving complementary feature fusion through late-layer concatenation.

CNN uses vibration sensor signals as inputs, while DNN uses nine time-domain statistical features of bearing vibration sensor signals as inputs [5]. The above diagnosis of bearing faults using convolutional neural networks have achieved good results.

At present, domestic and foreign research in this area is more and more extensive, but there is still the problem of lower recognition accuracy or longer fault recognition time, so this paper for the current automated production process of bearing fault diagnosis needs, from the bearing of the outer ring damage, the inner ring damage, rolling body damage and bearing good, the four aspects of the dataset of fault signals in the preprocessing of the labeling, based on Tensorflow deep learning framework, the respective fault diagnosis model, and based on Convolutional Neural Network(CNN) and Long Short-Term Memory Network (LSTM) after model training and data processing, the is finally achieved requirement of visual classification of fault types of rolling bearings. The innovation of this paper is to design a two-channel model combining convolutional neural network (CNN) and long-short-term memory network (LSTM) to improve the accuracy of fault identification and classification, and to optimize some previous methods.

2 Methodology

2.1 CNN model

CNN networks have end-to-end characteristics, which is capable of encompassing the entire pipeline of discriminative feature learning, feature dimensionality reduction and classifier classification by a single neural network [6]. Therefore, we use the just saved dataset to train the fault type classification network, vectorize the selected sample fault data as feature vector input to our designed WDCNN (CNN-based fault diagnosis) model, the model after data dimensionality reduction to identify and classify the fault types, and finally output the classification results [7].

This time we choose one-dimensional convolutional neural network (1DCNN) for our study. And the WDCNN (CNN-based fault diagnosis) model designed in this paper has two key points in terms of network construction:

The first layer of WDCNN uses a larger 64×1 convolutional kernel, which is designed to capture features over a short time range, a design idea similar to that of the Short-Time Fourier Transform (STFT), which uses a sinusoidal function as a window. However, the difference is that the large convolutional kernel obtained by the WDCNN through training optimization demonstrates self-optimizing feature discovery through discriminative learning mechanisms, effectively suppressing non-diagnostic signal components.

With the purpose of enhancing the performance of WDCNN, all the convolutional layers except the first layer use smaller convolutional kernels of 3×1. The network is deepened by reducing the number of convolutional kernels. After each convolution operation, adding the Batch Normalization (BN) operation to mitigate the phenomenon that the data input to the network generates bias to the pooling layer, so as to improve the efficiency of the model training and enhance the generalization ability of the model. Based on the above two points, a network structure diagram can be constructed and the network structure parameters can be determined.

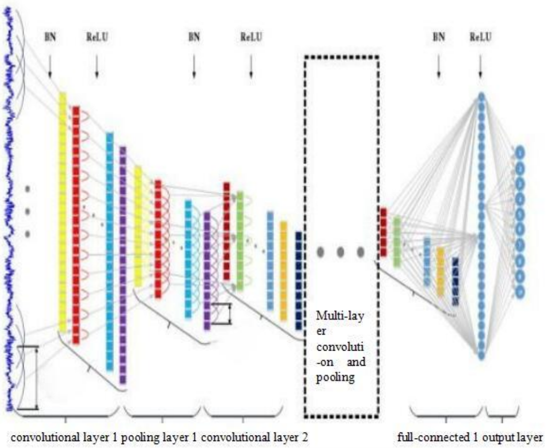

Fig. 1. CNN network architecture. (Photo credit: Original)

2.2 LSTM model construction

When LSTM processes ordered data from bearing vibration signals, first, bearing vibration data requires acquisition and preprocessing into tensor representations compatible with deep learning architectures. This usually includes signal sampling, normalization, and even feature engineering such as time domain to frequency domain conversion. Next, the processed signal is divided into serialized chunks of data. For example, a continuous vibration signal is segmented into a series of small segments, each containing signal data over a fixed length of time, serves as the input feed for the LSTM architecture [8].

The LSTM processes ordered data, especially bearing vibration signals, by controlling the flow of information through three main gates within it. The network architecture is as follows:

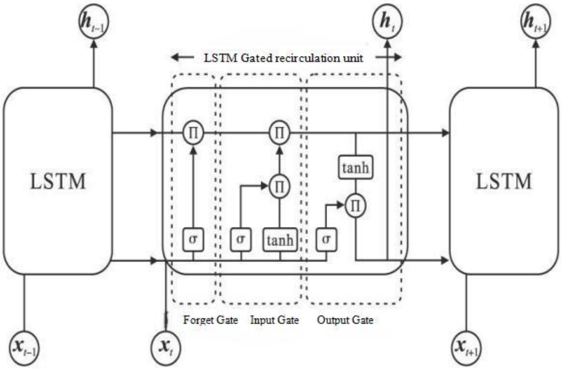

Fig. 2. LSTM network architecture. (Photo credit: Original)

Forget Gate: Determines what information needs to be forgotten or discarded from the unit state. This is especially important for ignoring non-important features, such as background noise, in bearing fault detection.

Input Layer (Input Gate): determines what information will be fed into the unit state. This allows the network to learn to recognize failure modes based on current steps and past information.

Output Layer (Output Gate): determines what information will constitute the output of the current step based on the current unit state. This determines how to recognize specific fault states based on current and past learning.

In this way, the LSTM architecture effectively captures sequential dependencies in bearing vibration waveforms through its recurrent processing mechanism. and memorize the failure modes for various time spans, and process the input vibration signal sequence step by step in a chronological order, processing one data block at each step and updating its internal state, that models sequential relationships within time-series data.

Based on this above mechanism, LSTM can learn the long-term patterns and short-term fluctuations during the bearing operation, and effectively recognize the complex patterns related to the failure. And at the end of the sequence, LSTM can give a prediction of the fault state of entire vibration waveform signal, synthesized through temporal dependencies extracted from the entire signal sequence.

2.3 The proposed CNN- LSTM model

As mentioned above, CNN can quickly and accurately extract the spatial features of the original vibration data, while LSTM can extract that cannot be recognized by

CNN the temporal information features of, so it is possible to combine the advantages of each of the two to design a two-channel fault diagnosis model to the raw vibration streams undergo feature fusion with the extracted feature information, and then finally to identify and classify the fused features [9]. In this way, the original CNN-based fault diagnosis model can be optimized, and the more complex patterns in the bearing vibration signals can be learned efficiently and accurately classified or regressed to identify potential fault types. This approach is capable of demonstrating high efficiency and accuracy .in monitoring and diagnosing the health status of rolling bearings, a time-dependent signal analysis task.

The spatial features extracted by the CNN layer can be further processed by the LSTM layer in order to integrate the features on the time axis while preserving the temporal dependency. And after the CNN and LSTM, there will be one or more fully connected layers and output layers which are used to map the extracted fusion features to fault types. By using on the final output of the network softmax activation function, the model can classify different types of bearing faults.

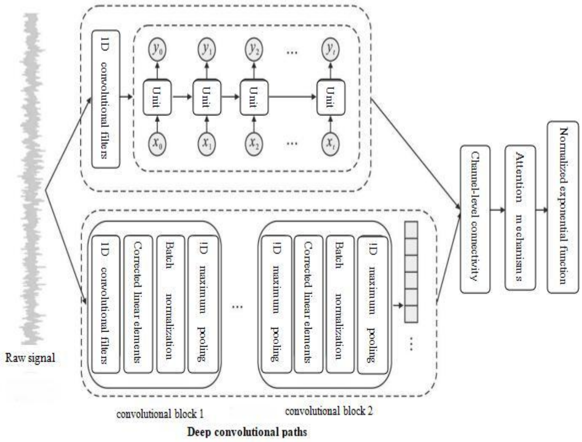

Fig. 3. CNN-LSTM dual path rolling bearing fault diagnosis model structure. (Photo credit: Original)

The deep convolution illustrated path in Figure 3 consists of five convolutional for extracting features stages and one for shrinking dimensionality reduction stage feature matrix. the extracted In the deep convolution path, a one-dimensional (1D) convolution filter, which is responsible for extracting the feature vectors required by the model from the input data; a linear unit that is corrected by the ReLU activation function; a batch normalization operation to optimize the training process; and a maximal pooling unit that improves translational invariance .with respect to the feature vectors that are input to the model.

This model is better at compared to the CNN model alonecapturing long range dependencies as well as extracting deeper vibration signal features. A 1D convolutional layer is added to the connected long short-term memory (LSTM) paths to further filter out high-frequency noise and enhance feature learning, providing optimized inputs for subsequent processing. An attention mechanism module is implemented to select discriminative features and dynamically reassign their weights, with the refined feature vectors then being transmitted to the fully connected layer for final classification. Eventually, the fused feature embeddings undergo softmax normalization to generate probabilistic outputs, enhancing classification precision.

3 Data set and experimental procedure

3.1 Data set

The dataset used in this study is provided by Case Western Reserve University (CWRU) Rolling Bearing Data Center. The object of is the figure below the experiment drive end bearing located in the left part of the experimental platform in. The diagnosed bearings have rolling body damage, outer ring damage and inner ring damage 3 kinds of failure conditions, the size of the damage diameter were including 0.007 inches, 0.014 inches and 0.021 inches, a total of 9 kinds of failure state and normal state, a total of 10 kinds of working conditions [10].

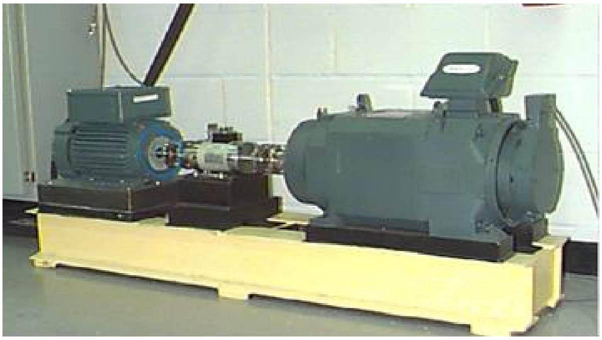

Fig. 4. Case Western Reserve University Bearing Experimental Platform错误!未找到引用源。

In this experiment, we counted the failure data under four loads, 0hp (horse power), 1hp, 2hp and 3hp, in the samples of this experiment. The 12k drive-end bearing failure data was selected with a sampling frequency of 12 kHz and the data was collected at 12,000 times/second [11].

Only DE_time is obtained for each type of dataset(drive-side acceleration) data.

The read data is sliced, i.e., the training data is divided into training data and test data. Divide the test set into test set and validation set so that the data is divided into three parts: training set, test set and validation set. After making labels for the training set besides test set, save the data.

3.2 Experimental procedure

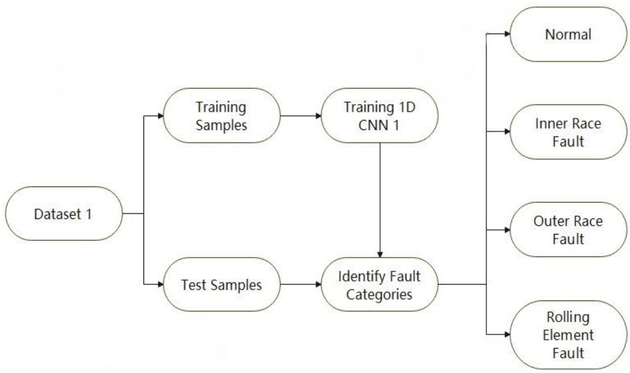

Fig. 5. The specific process of CNN fault diagnosis. (Photo credit: Original)

In bearing fault diagnosis, it is first necessary to gather the vibrational data during bearing operation, and according to the actual experimental needs of the collected raw data to cut and label, pointing out the different operating states or fault type. Then the time series data need to be preprocessed as necessary. The preprocessed data are divided into training set, validation set and test set, which are used for model learning, adjusting hyper-parameters and evaluating model performance, respectively, and are saved after labeling.

The output layer is designed as well as the appropriate activation function (e.g. softmax) is selected according to the needs of the classification task. After each epoch, model performance is enhanced through optimization of network topology, training cadence, batch dimensionality, and epoch progression, complemented by regularization strategies.

After training, the test set is imported into the model to complete the evaluation of the trained model. The model recognition accuracy and validation loss rate are recorded and analyzed, as well as the confusion matrix of the model recognition results are plotted, etc. to ensure the generalization of the model. The trained model needs to be saved and can later be loaded and deployed in real applications for real-time bearing failure monitoring.

Firstly, Input vectors are extracted from the raw dataset using 784-point temporal segments for model initialization, because the length of the samples affects the running time of the whole network and the accuracy of fault detection, so they should not be too long or too short. Each diagnostic class contains 500 data instances for model training.

The fault data in three sizes of 0.007 inches, 0.014 inches, and 0.021 inches under four loads of 0hp, 1hp, 2hp, and 3hp as well as the data for the normal condition are input to the model, and the input data are visualized after the t-sne dimensionality reduction function, as shown in the following figure:

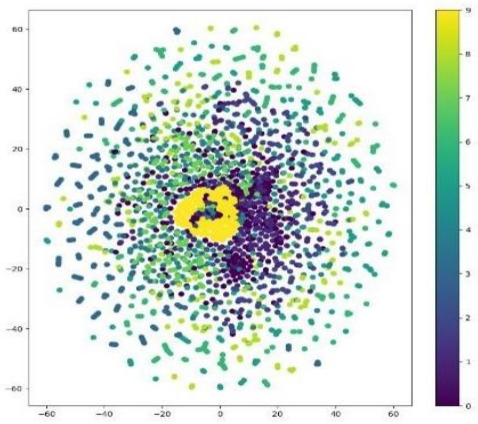

Fig. 6. Visualizing Input Data. (Photo credit: Original)

As can be seen from the above figure, the fault data is very cluttered and needs to be identified and categorized by the model, and then output the categorized fault type data. The normalized dataset volume reaches 1,250 samples, allocated through a 5:2.5:2.5 distribution among training, validation, and testing partitions with randomized shuffling. After disrupting the order, the vectorization operation of data features is performed and then input into the compiled model for identification and classification.

3.2.1 CNN fault diagnosis.

The model hyperparameters were adjusted so that batch_size (the number of samples captured in a training session) was 128 and the number of iterative training sessions was determined to be 200.

The previous fault data is fed into the trained model so that the model performs recognition and classification, the recognition accuracy is recorded and the recognition accuracy curve and the verification loss value curve are plotted as follows:

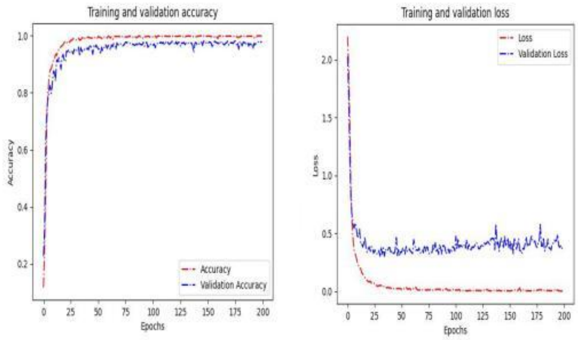

Fig. 7. Accuracy Curve and Lost Value Curve. (Photo credit: Original)

From Figure 7, it demonstrates that the validation loss value curve is not smooth compared to the training loss value curve, which is not a good match, so it can be concluded that there is an overfitting phenomenon in the training process of the model, which may identify and classify the noise in the fault data as well as some invalid data as well.

The classified data is again downscaled by the t-sne downscaling function to output the classified visualization results.

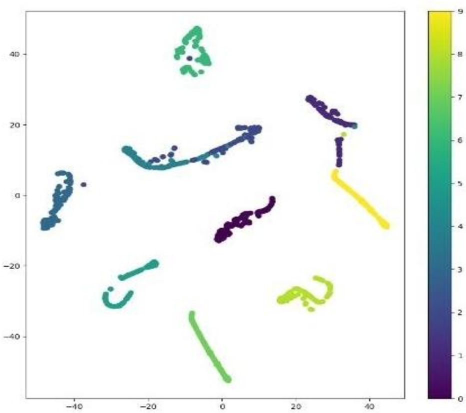

Fig. 8. Fault type classification results. (Photo credit: Original)

The above figure after classification is more concise and orderly than the initial input data, and the types of faults can be clearly understood through the classification of ten colors. However, there is still a part of overlapping, which indicates that non-sanitized data remnants persisted through the initial preparation stage, the deep learning architecture demonstrated remarkable generalization capability achieving ~95% accuracy. Figure 9 shows the confusion matrix.

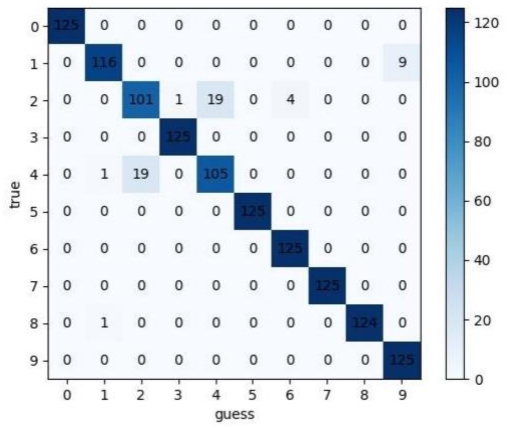

Fig. 9. Result of the confusion matrix. (Photo credit: Original)

3.2.2 LSTM fault diagnosis.

The model hyperparameters were adjusted so that batch_size (number of samples grabbed in one training session) was 128 and the number of iterative training sessions was determined to be 200.

The previous fault data is fed into the trained model so that the model performs recognition and classification, the recognition accuracy is recorded and the recognition accuracy curve and the verification loss value curve are plotted as follows:

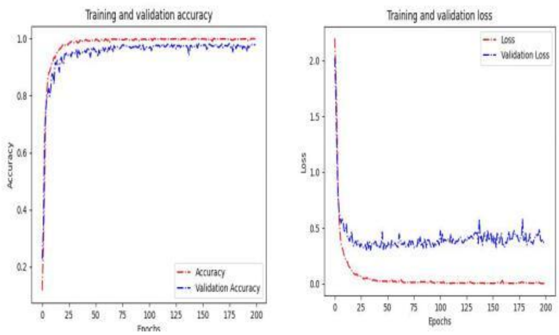

Fig. 10. Accuracy Curve and Lost Value Curve. (Photo credit: Original)

Because the separate LSTM algorithm model requires plenty of samples for training, when the number of samples is controlled to be the same as the CNN the LSTM model suffers from overfitting phenomenon during the training process, which indicates that its generalization ability is poor and the recognition time is longer than that of the CNN model.

The classified data is again downscaled by the t-sne downscaling function to output the classified visualization results.

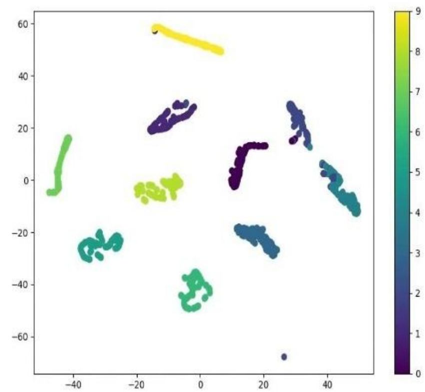

Fig. 11. Fault type classification results. (Photo credit: Original)

The above figure gives the fault data after detection and classification. The classification of ten colors gives a clear picture of the type of fault.

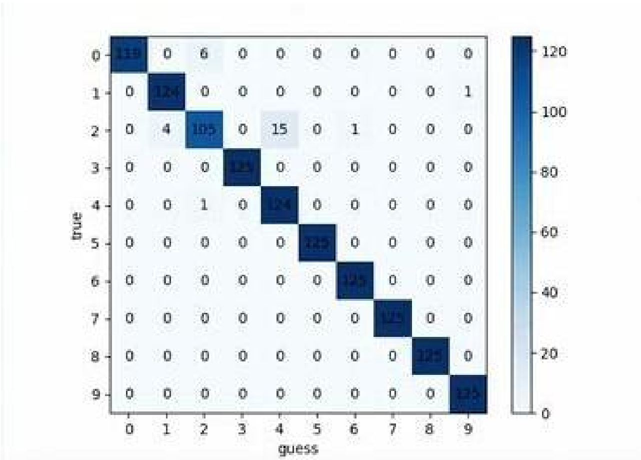

Fig. 12. Result of the confusion matrix. (Photo credit: Original)

From the confusion matrix in figure 12, we can see that the recognition accuracy is basically around 96%.

3.2.3 CNN-LSTM fault diagnosis.

The model hyperparameters were adjusted so that batch_size (number of samples grabbed in one training session) was 128 and the number of iterative training sessions was determined to be 400.

The fault data that has been vectorized earlier is fed into the trained model so that the model performs recognition and classification, the recognition accuracy is recorded and the recognition accuracy curve and the verification loss value curve are plotted as follows:

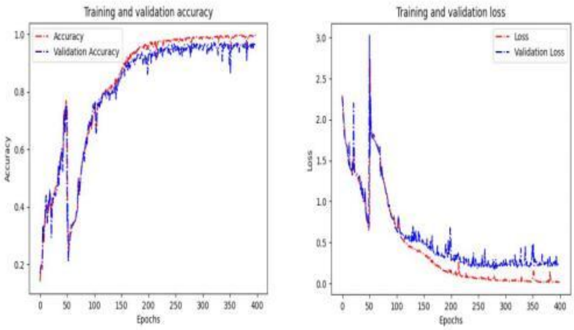

Fig. 13. Accuracy Curve and Lost Value Curve. (Photo credit: Original)

From Figure 7, it demonstrates that the validation loss value curve coincides with the training loss value curve, indicating that there is no overfitting phenomenon during the model training process, and therefore the CNN-LSTM model has a very good generalization ability.

The classified data is again downscaled by the t-sne dimensionality reduction function, and the output classified visualization results are shown in the following figure:

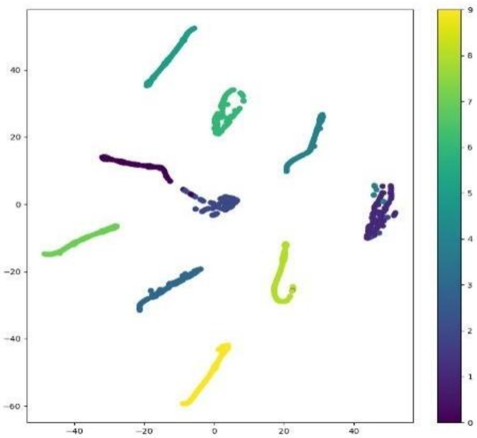

Fig. 14. Fault type classification results. (Photo credit: Original)

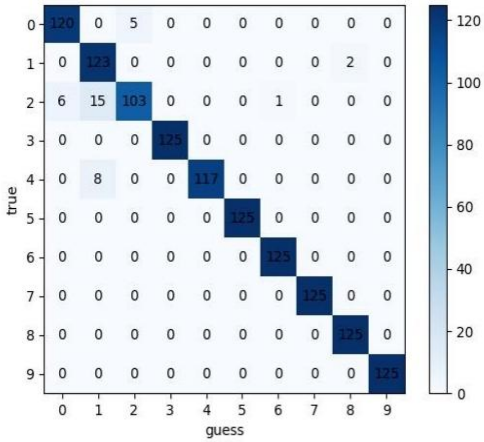

Fig. 15. Result of the confusion matrix. (Photo credit: Original)

From these above two figures, we can visually observe that the classified fault types are more concise and organized than the initial input data, with a very small number of overlapping parts, and the recognition accuracy is basically consistent with the "98% or so" result described earlier.

4 Comparison summary

CNN, LSTM and CNN-LSTM have their own advantages and disadvantages in different scenarios and applications. They show their respective advantages in processing image data (e.g. CNN) and sequence data (e.g. CNN-LSTM).

Table 1. Comparison of the fault diagnosis perfomance of different methods

Methodology	Recognition accuracy	Time required
CNN	95%	5-7min
LSTM	96%	10-12min
CNN-LSTM	98%	13-15min

CNN has the ability to inherently learn discriminative representations from raw vibrational waveform data and facilitate spatial-temporal feature extraction in chronological datasets. On the other hand, LSTM is capable of handling long-term temporal dependencies. Therefore, combining CNN with LSTM allows for capturing spatial features (via CNN) as well as temporal relationships (via LSTM) in bearing vibration signals with sequential logic. Due to the memory capability of LSTM, CNN-LSTM model is more suitable for handling complex time series data, as it can remember and learn patterns and dependencies in long sequences, offering better generalization capabilities compared to single models.

The main drawbacks of CNN-LSTM lie in computational cost and network complexity. When performing fault signal recognition and analysis using the TensorFlow deep learning framework, the CNN-LSTM model steadily improved recognition accuracy but took significantly longer than the two single models. This is because LSTM has a recursive nature, with a large number of model parameters, requiring more computational resources during training, which can result in longer training times compared to using CNN alone. Additionally, due to the more complex model structure compared to standalone CNN, more careful tuning of hyperparameters and adjustments to the network architecture are needed to achieve optimal performance.

In summary, standalone CNN excels in feature extraction and computational efficiency, making them suitable for images and other high-dimensional data. LSTM, on the other hand, is capable of capturing temporal relationships in bearing vibration signals with sequential logic and improving generalization capabilities. The CNN-LSTM leverages the strengths of convolutional neural networks and recurrent neural networks simultaneously, enabling the extraction of spatial features while capturing deeper temporal relationships and features in sequential logic signals, at the cost of requiring more computational resources and meticulous model tuning.

5 Conclusion

This paper focuses on the rolling bearing in mechanical equipment as the research object, for the phenomenon of identifying and diagnosing faults and then proceeding to repair them, which is common in today's industry, we start from the aspect of fault detection of the vibration signals of rolling bearings, and research on the practical application of deep learning in this aspect. On the basis of existing research, faults in rolling bearing vibration signals are identified and classified by using CNN-based and LSTM-based fault diagnosis models. This project makes innovations in CNN model network construction and embedding LSTM algorithms in it, and proposes a dual-channel fault diagnosis model based on CNN-LSTM. The study was conducted using the bearing vibration signal dataset from Case Western Reserve University, U.S.A. The fault signals were identified and classified by the fault diagnostic models that have been designed, and the respective fault identification accuracy and visualization classification results of each model were obtained. After designing the new model, the fault recognition accuracy was improved from 95% to 97%. It is hoped that subsequent research can further enhance data standardization, optimize the required time, and reduce time costs, thereby bringing greater convenience to industrial production.

References

1. Zheng, F.: 'Research on Mechanical Equipment Fault Early Warning and Diagnosis Technology Based on Deep Learning'. Master thesis, Beijing University of Chemical Technology, 2020.
2. Zhang, Z.: 'Deep learning and its application in health monitoring and fault diagnosis of mechanical equipment'. Master thesis, Beijing University of Chemical Technology, 2022.
3. Wang, Y.: 'Research on bearing fault diagnosis method based on deep learning'. Master thesis, Beijing University of Posts and Telecommunications, 2023.
4. Chen, Z., Mauricio, A., Li, W., Gryllias, K.: 'A deep learning method for bearing fault diagnosis based on Cyclic Spectral Coherence and Convolutional Neural Networks', Mechanical systems and signal processing., 2020, 140, (1), pp. 106683.
5. Li, H., Huang, J., Ji, S.: 'Bearing Fault Diagnosis with a Feature Fusi on Method Based on an Ensemble Convolutional Neural Network and Deep Neural Network', Sensors., 2019, 19, (9), pp. 2034.
6. Wang, X., Li, A., Han, G.: 'A Deep-Learning-Based Fault Diagnosis Method of Industrial Bearings Using Multi-Source Information', Applied Sciences., 2023, 13, (2), pp. 933.
7. Wang, Q., Deng, L., Zhao, R.: 'Rolling bearing fault diagnosis based on one-dimensional CNN migration learning', Journal of Vibration Measurement and Diagnosis., 2023, 43, (1), pp. 24-30.
8. Tang, H., Li, B., Gao, S.: 'A CNN-LSTM bearing fault diagnosis method based on dual-channel feature fusion', Digital Manufacturing Science., 2022, 20, (4), pp. 253-257.
9. Jing, S., Wu, D. 'LSTM-CNN based dual path rolling bearing fault diagnosis', Journal of Shenyang University of Technology., 2024, 1, (1), pp. 44-49.
10. Wang, X., Hu, T., Xi S.: 'Bearing fault diagnosis method based on convolutional neural network', Forestry Machinery and Woodworking Equipment., 2021, 49, (10), pp. 50-59
11. Hei, J.: Research of Bearing Fault Diagnosis System Based on Edge Computing'. Master thesis, Xidian University, 2022.