Identification and Prediction of Magnetorheological Damper Force Based on Deep Learning

Wei Huang

SINOMACH Academy of Science and Technology Co. Ltd, SINOMACH Research Center of Engineering Vibration Control Technology, Beijing 100080, China.
Key Laboratory of Vibration Control for Key Process Equipment of Major Scientific Infrastructure in the Machinery Industry, Beijing 100080, China. E-mail: huangweiac@126.com

Jian Xu

China National Machinery Industry Corporation Ltd (SINOMACH), Beijing 100080, China.

(Received 18 March 2025; accepted 12 August 2025)

Magnetorheological dampers (MRDs) are widely used in the field of engineering vibration control. The accurate identification and prediction of their output forces are crucial for optimizing control strategies. However, traditional analysis methods based on mechanical models and empirical formulas have many limitations. This study proposes an innovative deep-learning approach. First, the continuous wavelet transform (CWT) is employed to convert the one-dimensional signals of MRD output force into two-dimensional time-frequency maps. Then, a convolutional neural network (CNN) is employed for feature extraction and type identification, and a CWT-CNN model is constructed. This model achieves 100% accuracy on the test dataset. In addition, by combining the local feature extraction ability of CNN and the sequence modeling advantage of the long short-term memory network (LSTM), a CNN-LSTM model is built to predict the MRD output force. The results show that compared with CNN and LSTM, the CNN-LSTM model exhibits stronger generalization ability. It outperforms the other models in comprehensive performance, as evaluated by MSE, RMSE, MAE, MAPE, and R^2. This study provides an effective technical means for the identification and prediction of MRD output forces and promotes the application and development of deep learning in this field.

NOMENCLATURE

MRDs:	Magnetorheological dampers
CWT:	Continuous wavelet transform
CNN:	Convolutional neural network
LSTM:	Long short-term memory network
RNN:	Recurrent neural network
MSE:	Mean Squared Error
RMSE:	Root Mean Squared Error
MAE:	Mean Absolute Error
MAPE:	Mean Absolute Percentage Error
R^2:	Coefficient of determination
T-SNE:	T-Distributed Stochastic Neighbor Embedding
AP:	Average pooling
MP:	Max pooling

1. INTRODUCTION

Magnetorheological dampers (MRDs), a key component in intelligent materials, exhibit excellent performance in engineering vibration control and are widely used in fields such as seismic protection of building structures, bridge vibration reduction, and vehicle suspension systems.[1–5]

The output characteristics of MRDs are influenced by multiple factors, including excitation amplitude, vibration frequency, and control current or voltage. Precise identification and prediction of an MRD force are crucial for optimizing control strategies and enhancing vibration control effectiveness. Traditional methods for analyzing MRD output, such as analytical methods based on mechanical models and empirical formula-based methods, have numerous limitations when dealing with complex working conditions. Mechanical models, often based on simplified assumptions, fail to accurately describe the nonlinear behavior of MRDs in practical applications. Empirical formulas, on the other hand, rely heavily on specific experimental conditions and lack universality. With the rapid development of deep learning technology, its powerful feature extraction and data modeling capabilities have provided new insights and methods for solving the problems of identifying and predicting MRD output.

In the realm of mechanical models and empirical formulas for MR materials, numerous scholars have conducted extensive and representative research, encompassing performance prediction, reverse design, parameter identification, and vibration control optimization. Liao et al.[6] addressed the uncertainty in the relationship between magnetic current and natural frequency of magnetorheological elastomer (MRE) dynamic vibration absorbers (DVAs) caused by material nonlinearity. They proposed a stiffness tuning algorithm based on phase difference, which can rapidly track the excitation frequency and exhibit high efficiency in vibration control. Ren et al.[7] constructed a deep learning framework integrating generators and predictors for the rapid reverse design of MREs. By optimizing the loss function, they solved the problem of multiple solutions in reverse design, providing an efficient tool for the in-

dustrialization of MREs in fields such as vibration control and soft machinery. Furthermore, Ren et al.[8] proposed the MDCBA model, which combines magnetic dipole theory with a CNN-BiLSTM-Attention architecture, for predicting the magnetically induced shear storage modulus of MREs under small sample data. The transfer performance of the model under different silicone oil contents and loading frequencies verified its generalization ability, offering an efficient method for characterizing MRE performance. Guo et al.[9] investigated the influence of particle swarm optimization (PSO) algorithm parameters on the parameter identification accuracy of MRD models. By analyzing the optimal combination of parameters such as the number of particles, number of iterations, and learning factors, they achieved the highest accuracy in predicting the damping force of the hyperbolic tangent model. Bedoya-Zambrano et al.[10] proposed a novel controller combining the NSGA-II genetic algorithm with fuzzy logic for optimizing the control force of MRD in three-dimensional frame structures. Comparative studies validated its advantages in vibration control of complex structures. Lv et al.[11] tackled the issue of discrepancies between the rheological properties of magnetorheological fluids (MRFs) at high shear rates and the test data at low shear rates. They proposed the exponential linear mixing analysis (ELMA) model, which can accurately predict the behavior of MRFs at extremely high shear rates.

Among the many deep-learning models, the convolutional neural network (CNN) can automatically extract local features from data when processing images and signals, achieving remarkable results in related fields.[12,13] However, when faced with the dynamically changing vibration signals of MRDs, CNN has limitations in capturing long-term dependencies in the signals. The long short-term memory network (LSTM),[14,15] an optimized version of the recurrent neural network (RNN),[16,17] has successfully overcome the problems of gradient vanishing and explosion through its innovative gating mechanism. When processing time-series data, LSTM demonstrates excellent long-term information-memory capabilities. Nevertheless, LSTM is relatively weak in local feature extraction.

In view of this, this study innovatively proposes a phased and comprehensive research plan. First, the continuous wavelet transform (CWT) is utilized to deeply process the output force of MRD, converting one-dimensional signals into two-dimensional time-frequency diagrams that contain rich time-frequency information. This conversion not only expands the feature dimensions of the signals but also provides more valuable input for the subsequent CNN model. Based on these time - frequency diagrams, the CNN can efficiently extract the key features of MRD output, enabling precise identification of MRD output. On this solid foundation, a CNN-LSTM model is further constructed. By fully integrating the powerful feature-extraction ability of CNN and the excellent time-series modeling advantage of LSTM, the CNN-LSTM model is used to predict the MRD output. Through a systematic comparison of the performance of CNN, LSTM, and CNN-LSTM models in MRD output identification and prediction tasks, this study aims to provide a scientific and accurate basis for model selection in engineering practice, and strongly promote the in-depth application and development of deep-learning technology in the field of MRD vibration control.

2. MECHANICAL TEST DATA OF THE MRD

Experiments were conducted to study the dynamic characteristics of a certain model of MRD produced by Lord Corporation in the United States. The excitation signal was a sine wave, with excitation amplitudes set at 5 mm, 10 mm, and 15 mm respectively, and the vibration frequencies were set at 0.5 Hz, 1.0 Hz, 2.0 Hz, and 3.0 Hz, and the control currents of the MRD were set at 0 A, 0.1 A, 0.2 A, 0.3 A, 0.4 A, 0.5 A, 0.6 A, 0.7 A, 0.8 A, 0.9 A, and 1.0 A. Figure 1 and 2 show the mechanical characteristics of the MRD under three excitation amplitudes, four excitation frequencies, and eleven control currents.

Figure 3 presents the full-sample time series of the MRD output force. The data are arranged as follows: at a frequency of 0.5 Hz, with amplitudes of $\pm$5 mm/$\pm$10 mm/$\pm$15 mm, the output forces of the MRD with input currents ranging from 0 A to 1 A are given. Similarly, at a frequency of 1.0 Hz, for amplitudes of $\pm$5 mm/$\pm$10 mm/$\pm$15 mm, the corresponding MRD output forces with input currents from 0 A to 1 A are included. The same pattern applies to the cases at 2.0 Hz and 3.0 Hz frequencies, where for amplitudes of $\pm$5 mm/$\pm$10 mm/$\pm$15 mm, the MRD output forces with input currents from 0 A to 1 A are presented respectively.

3. IDENTIFICATION OF MRD FORCE

In the realm of engineering applications involving MRDs, accurately identifying their output is of great significance. This section delves into the methods and processes for precisely determining the output of MRDs, aiming to provide a comprehensive understanding and effective approach to this crucial aspect.

3.1. CWT

CWT is a time-frequency analysis method designed for time-varying and non-stationary signals.[18,19] It can transform one-dimensional vibration signals into two-dimensional time-frequency depictions, which encompass information from both the time and frequency domains.

Unlike the short-time Fourier transform, which uses a fixed window function, the continuous wavelet transform employs an adjustable window function. This function can achieve a balance between time and frequency resolution when analyzing non-stationary signals.

For an arbitrary signal $x(t)$, its continuous wavelet transform is defined as:

$$X_\omega(u,v) = \int_{-\infty}^{+\infty} x(t)\varphi_{u,v}(t)dt = \frac{1}{\sqrt{v}} \cdot \int_{-\infty}^{+\infty} x(t)\varphi\left(\frac{t-u}{v}\right)dt. \quad (1)$$

Here, u represents the translation factor, which decides the position of the wavelet window in the time domain. v is the scaling factor, which modifies the size of the wavelet window and

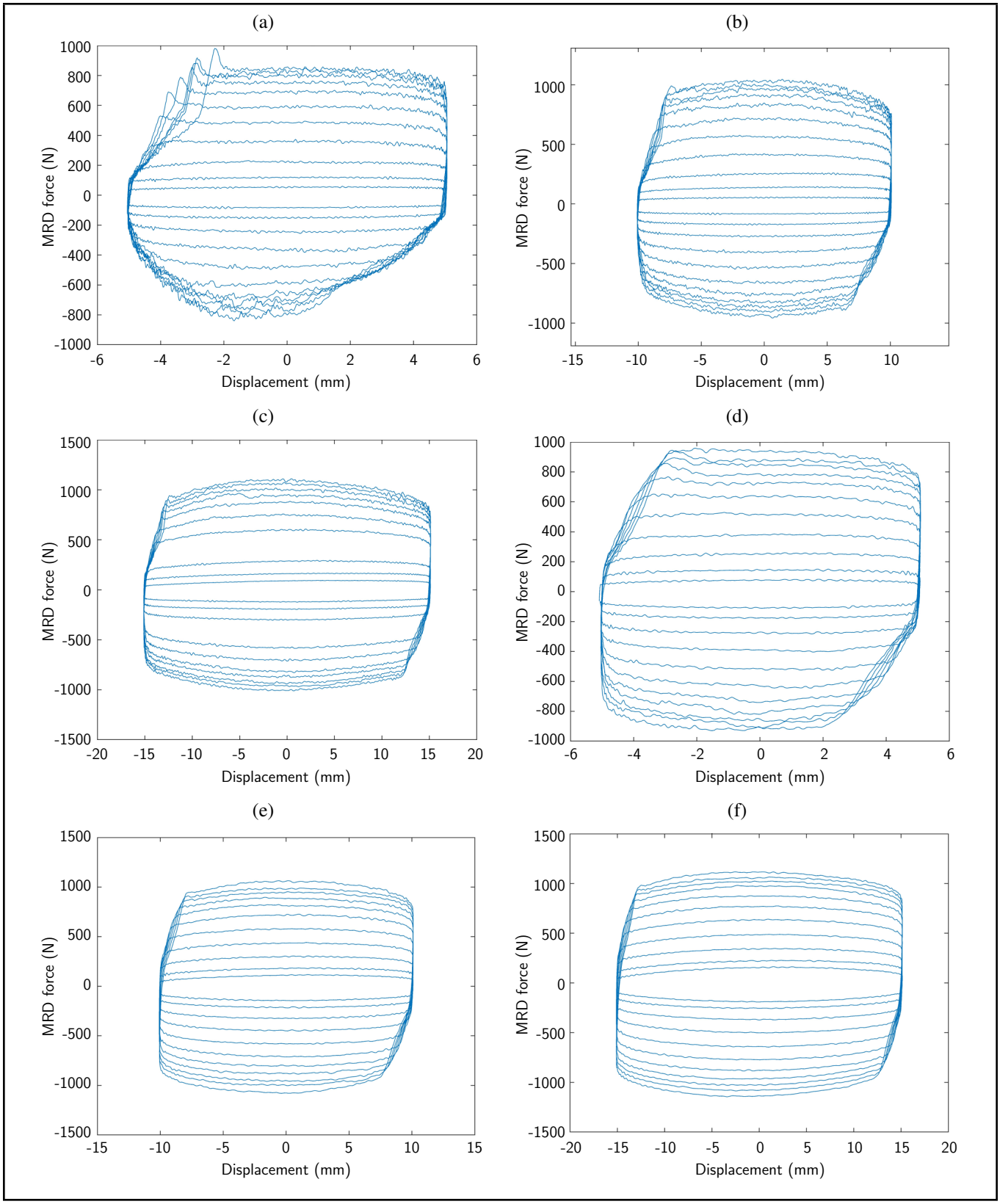

Figure 1. Experimental results of the mechanical characteristics of the MRD (part one). (a) Frequency = 0.5 Hz, amplitude = 5 mm; (b) Frequency = 0.5 Hz, amplitude = 10 mm; (c) Frequency = 0.5 Hz, amplitude = 15 mm; (d) Frequency = 1.0 Hz, amplitude = 5 mm; (e) Frequency = 1.0 Hz, amplitude = 10 mm; (f) Frequency = 1.0 Hz, amplitude = 15 mm

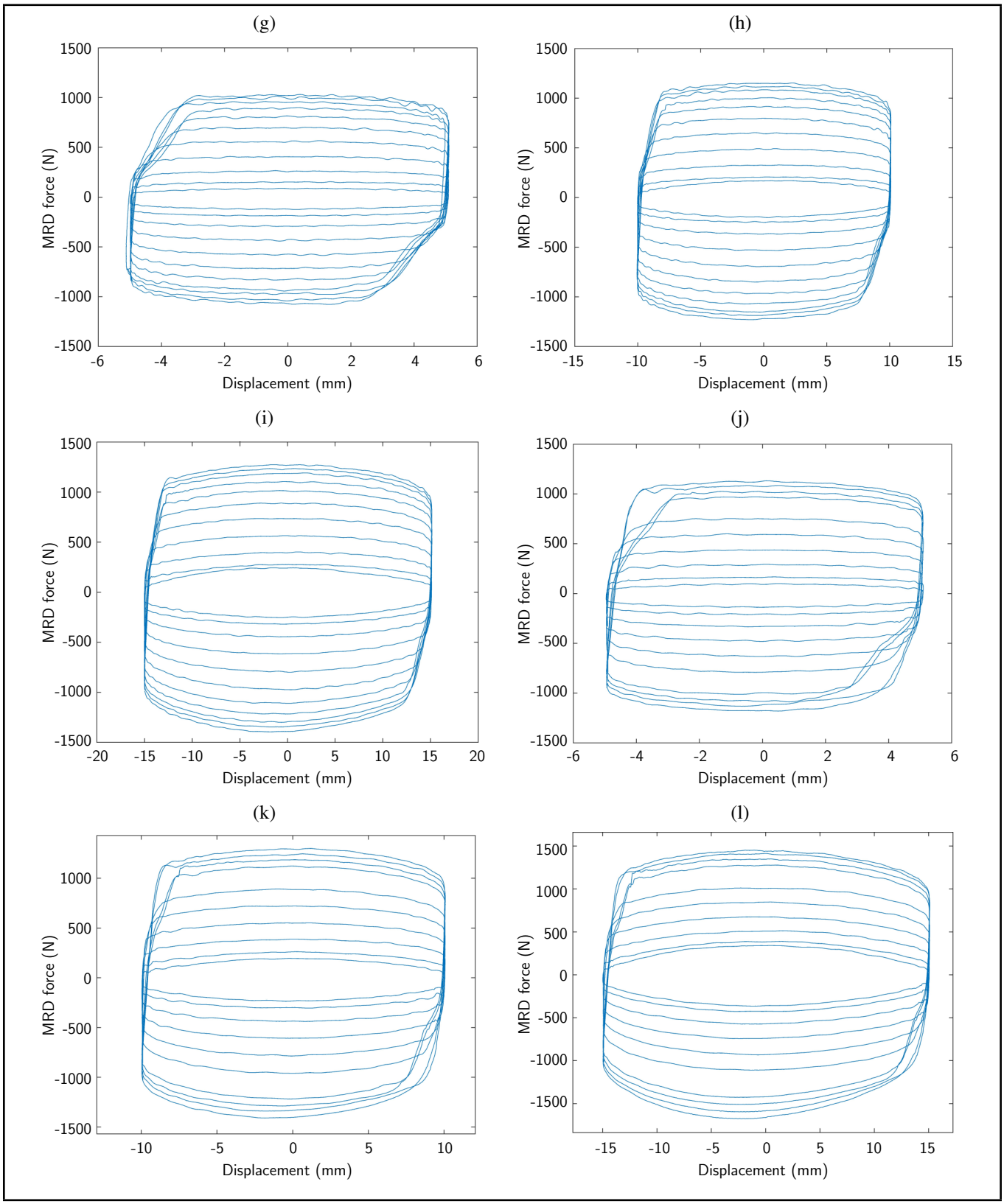

Figure 2. Experimental results of the mechanical characteristics of the MRD (part two). (g) Frequency = 2.0 Hz, amplitude = 5 mm; (h) Frequency = 2.0 Hz, amplitude = 10 mm; (i) Frequency = 2.0 Hz, amplitude = 15 mm; (j) Frequency = 3.0 Hz, amplitude = 5 mm; (k) Frequency = 3.0 Hz, amplitude = 10 mm; (l) Frequency = 3.0 Hz, amplitude = 15 mm

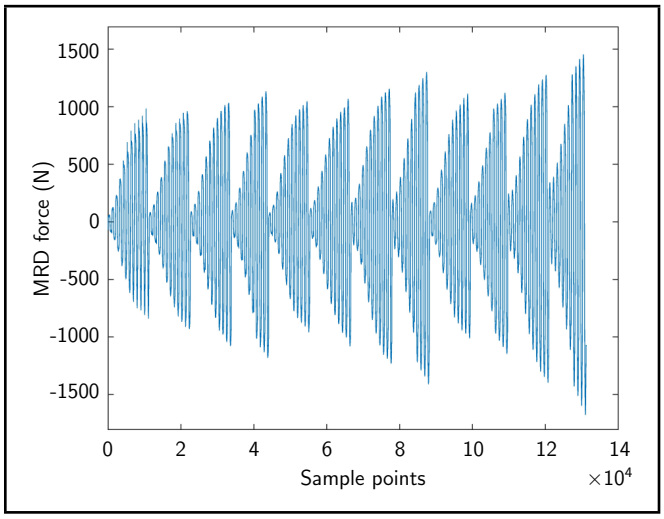

Figure 3. Full-sample time series of the MRD output force.

its position in the frequency domain. $\varphi_{u,v}(t)$ is the wavelet basis function (also known as the parent wavelet), and it is expressed as:

$$\varphi_{u,v}(t)dt = \frac{1}{\sqrt{|u|}}\varphi\left(\frac{t-v}{u}\right) \qquad u > 0. \tag{2}$$

In this research, the complex Morlet wavelet is chosen as the basis function for the continuous wavelet transform.

3.2. CNN

CNN is a type of feed-forward neural network that draws inspiration from the biological visual perception mechanism. It stands as one of the classic algorithms in the realm of deep learning.

The CNN model primarily consists of three core components: the convolutional layer, the pooling layer, and the fully connected layer (Fig. 4). What sets it notably apart from traditional neural networks is that in the convolutional layer, CNN replaces a matrix multiplication with convolution operations. It allows the local receptive field to scan across the entire dataset through sliding convolution, which substantially cuts down the number of parameters in the network. The specific convolution calculation formula is as follows:

$$x_i^j = f\left(\sum_{k=1} W_i^j * x_i^{j-1} + b_i^j\right). \tag{3}$$

Here, x_i^jdenotes the i-th feature of the output value of the j-th layer, W_i^j represents the weight matrix of the i-th convolution kernel of the j-th layer, x_i^{j-1} is the $(j-1)$-th output layer, and b_i^j is the bias term.

The pooling layer, also known as the down-sampling layer, is mainly utilized to reduce the parameters of the neural network. Commonly used pooling functions are average pooling (AP) and max pooling (MP).[20,21] Their mathematical expressions are as follows:

$$P_{AP}^{l(i,j)} = \frac{1}{W}\sum_{t=(j-1)W+1}^{jW} a^{l(i,t)}; \tag{4}$$

$$P_{MP}^{l(i,j)} = \max_{(j-1)W+1\le t\le jW}\left\{a^{l(i,t)}\right\} b_i. \tag{5}$$

In these equations, W represents the width of the pooling area, $p^{l(i,j)}$ is the activation value of the t-th neuron in the i-th frame of the l-th layer, and $a^{l(i,j)}$ represents the pooling result.

The fully connected layer is mainly for further feature extraction. The features obtained from the last pooling layer are flattened into one-dimensional feature vectors. These vectors are then fed into the classifier after undergoing non-linear activation by the ReLU function.

3.3. Proposed CWT-CNN

This research put forward a multi-source vibration recognition approach named CWT-CNN. It leverages the CWT to transform one-dimensional signals into two-dimensional time - frequency depictions, thus making it possible to represent vibration feature information. These two-dimensional time-frequency depictions are then fed into the CNN, which autonomously extracts relevant features. Eventually, the Softmax layer classifies the vibration type. The detailed procedures are as follows:

- The initially collected vibration data is randomly split according to the specified sample length.
- The CWT is employed to turn one-dimensional signals into two-dimensional time-frequency depictions.
- The obtained time-frequency depictions are proportionally separated into training and test datasets.
- A CNN model is constructed, and its parameters are initialized.
- The time-frequency depictions of the training dataset are input into the convolutional layers for model training, and the optimal model parameters are saved.
- The test dataset is input into the model for classification, and the classification outcomes and accuracy are assessed.

The architecture of the proposed CWT – CNN model in this research is presented in Fig. 5. First, the CWT technique is utilized to convert the one-dimensional vibration signal into a two - dimensional time-frequency image. Next, the two-dimensional time - frequency images are input into the convolutional layer of the CNN to automatically extract features. Finally, the vibration type is recognized via the SoftMax layer.

The CWT-CNN identification process for the MRD force is summarized in Fig. 6.

From 3 MRD force datasets, 60 samples are extracted for each type, resulting in a total of 180 samples. Each sample contains 1,024 sampling points, which are transformed into 64×64 sample graphs to act as model inputs. The dataset is then divided into training and test datasets at a ratio of 3:1.

T-Distributed Stochastic Neighbor Embedding (T-SNE) is a highly effective nonlinear dimensionality reduction technique. It is employed to map high-dimensional data into a

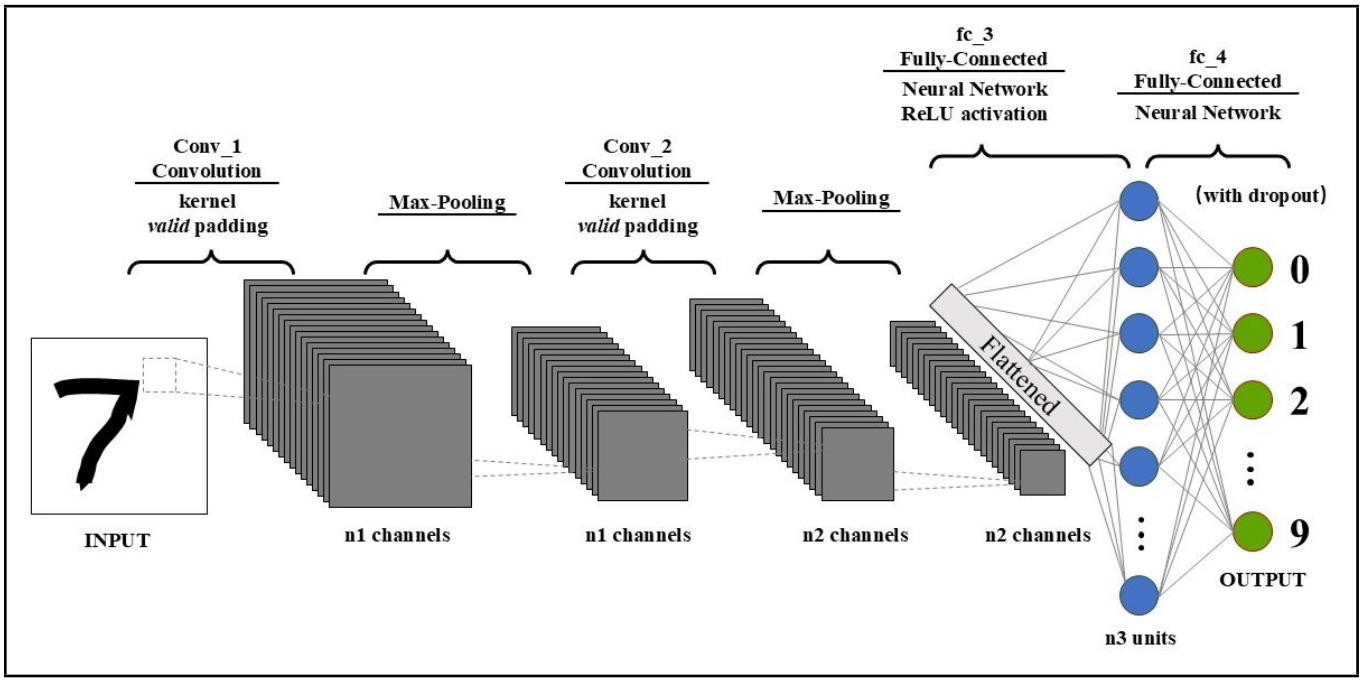

Figure 4. Schematic diagram of the CNN structure.

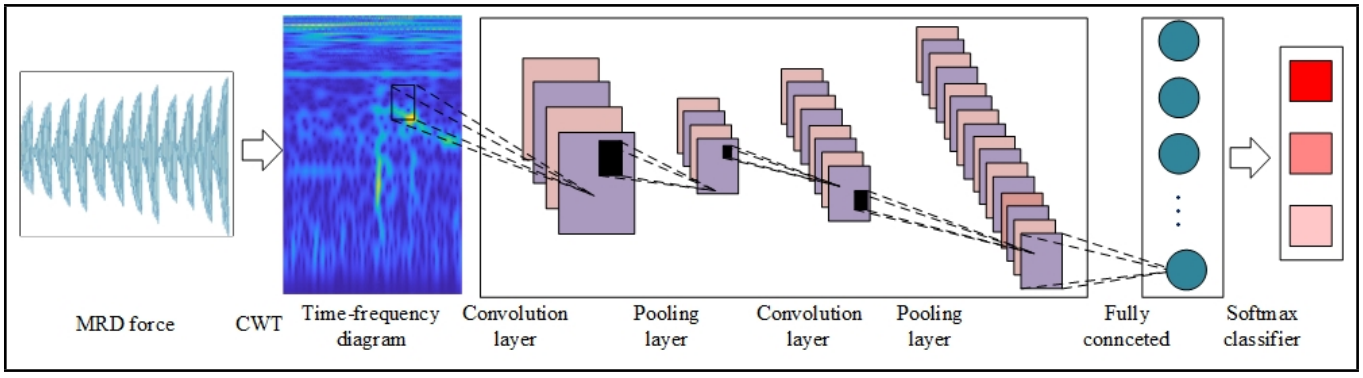

Figure 5. Basic structure of CWT-CNN model.

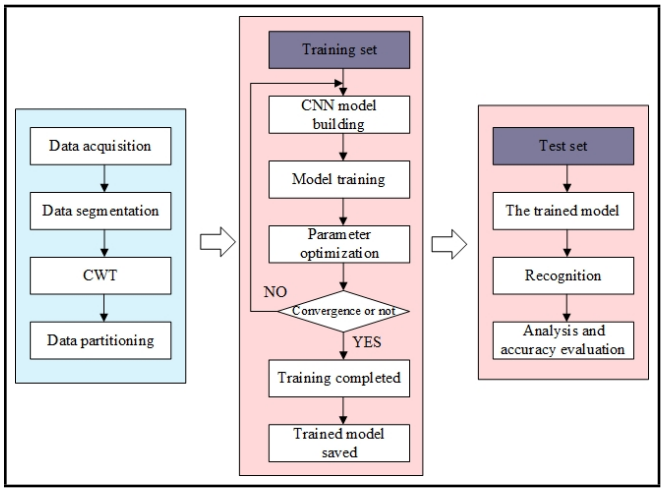

Figure 6. CWT-CNN identification process for the MRD force.

low-dimensional space, facilitating its visualization.[22,23] In order to visually assess the performance of the proposed model, T-SNE is utilized for visual analysis.

The visualization outcomes of the initial training set data are illustrated in Fig. 7. Here, 3 different colors are used to denote the 3 various inputs within the dataset (the three different MRD output inputs here refer to the classification based on different amplitudes: ± 5 mm, ± 10 mm, and ± 15 mm). The results demonstrate that the distribution of the original data are rather chaotic. There is a substantial overlap among the force signals from the 3 categories, rendering it infeasible to classify and identify MRD force solely based on the original data.

Figure 8 shows the visualization results of the data that have been processed by the CWT-CNN model. When compared with the original data, the processed data show a decreased level of disarray. They exhibit a tendency towards greater orderliness, and the features from different categories have become more distinguishable.

The classification performance of the model is depicted through a confusion matrix. As is evident from Fig. 9, the model attains an accuracy rate of 100% on the test dataset. This indicates its proficiency in accurately differentiating between various MRD force types.

4. PREDICTION OF MRD FORCE

In the field of engineering vibration control, accurately predicting the output of MRDs is a crucial task. Given the com-

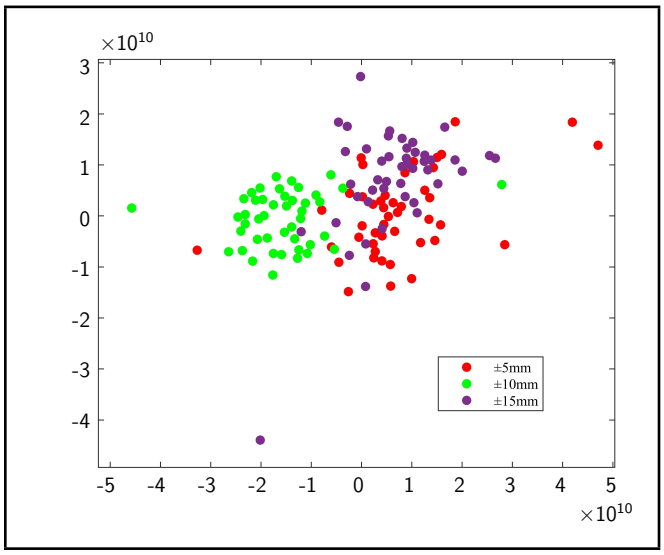

Figure 7. Original sample distribution.

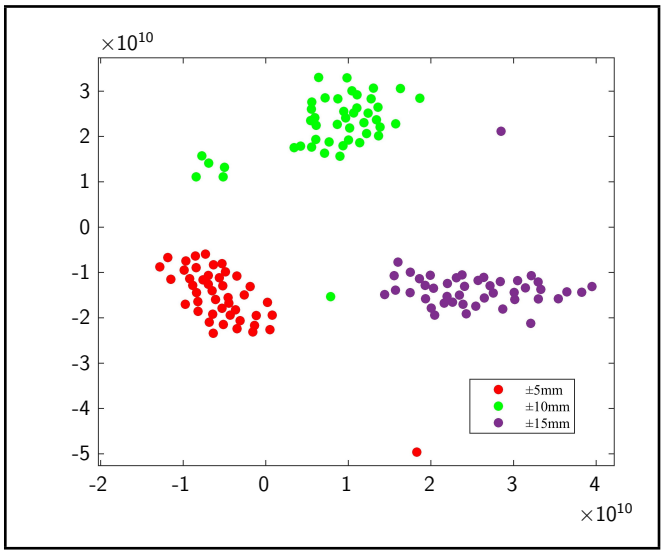

Figure 8. Sample distribution after model identification.

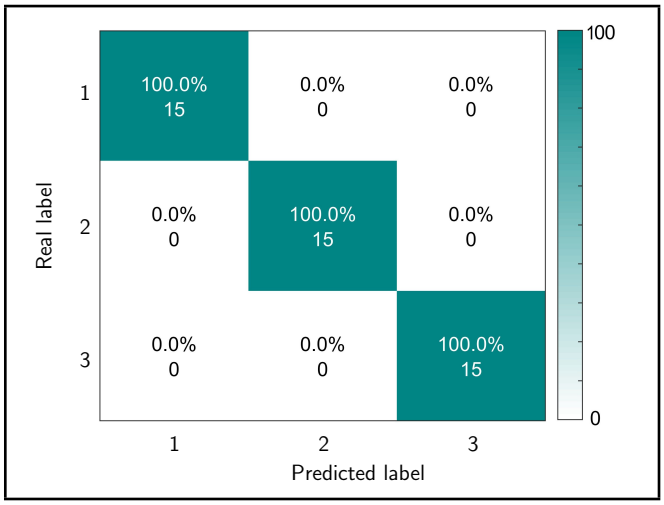

Figure 9. Confusion matrix for classification results. The accuracy rate at the test set is 100%.

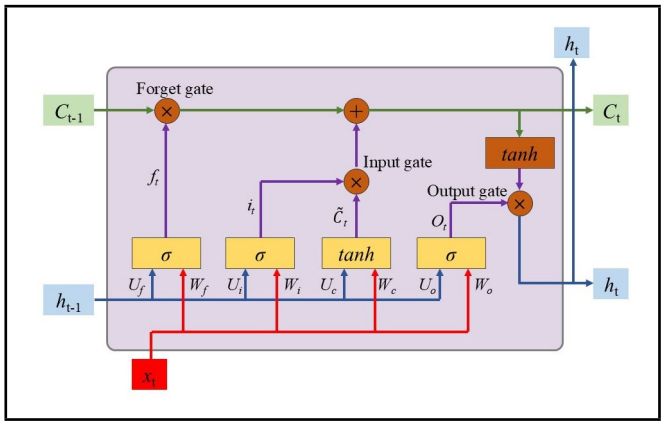

Figure 10. LSTM cell structure. [24]

plex and variable factors that affect MRD output, such as excitation amplitude, vibration frequency, and control current, precise prediction can help optimize control strategies and enhance the overall performance of vibration control systems. This section focuses on exploring and presenting methods for predicting the output force of MRDs, providing valuable insights and effective solutions for relevant engineering applications.

4.1. LSTM

For signals with time-scale periodicity, earlier signals can influence later signals. Nevertheless, traditional neural networks typically fail to take this influence into account. As a result, they often struggle to yield satisfactory outcomes when dealing with such signals.

To address this issue, scholars proposed the RNN. Inside the RNN, neurons form loops. When processing time-series signals, these loops can preserve earlier-time information for use in later-time calculations.[16,17]

Building on the RNN, scholars further developed the LSTM model. This model incorporates a gating mechanism to regulate the information state. It allows the network to retain significant information and discard unimportant information, thereby resolving the issue of gradient vanishing/explosion that might arise in the RNN model.

The schematic illustration of the LSTM unit is presented in Fig. 10. Its central part is the section where c_{t-1} is transferred to c_t t the top of the figure. This section is commonly referred to as the cell state, which persists throughout the entire LSTM chain system. It can store some long-term memory information and represents the key distinction between LSTM and traditional RNN.

Its calculation formula is as follows:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \widetilde{c}_t. \tag{6}$$

In the given formula, c_t and c_{t-1} represent the state vectors of the memory unit at time instants t and $t-1$ respectively. Their dimensions are set manually. The $\widetilde{c}_t$ is the updated value of the cell state at time t. The $f_t$and i_t are the forget gate and input gate at time t correspondingly.

All the elements in these two vectors range from 0 to 1. Specifically, f_t is in charge of discarding certain elements of

the state vector c_{t-1} from the previous time step, while i_t is responsible for introducing the effective elements from the current state update value $\widetilde{c}_t$. It's important to note that when multiplying the gate vector with the state vector, it's an element-wise multiplication, not matrix multiplication.

The computational approaches for f_t, i_t, and $\widetilde{c}_t$ are as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1},\ x_t] + b_f\right); \tag{7}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1},\ x_t] + b_i\right); \tag{8}$$

$$\widetilde{c}_t = \tanh\left(W_c \cdot [h_{t-1},\ x_t] + b_c\right); \tag{9}$$

where h_{t-1} is the output of the unit at the previous moment; x_t is the input of the unit at this moment. These two vectors are jointly used as the input for calculating the forget gate f_t, the input gate i_t , and the cell state update value $\widetilde{c}_t$; W_f, W_i, and W_c are the weights for calculating f_t, i_t, and $\widetilde{c}_t$, respectively; b_f, b_i, and b_c are the biases for calculating the above three variables. Since the elements in the forget gate f_t and the input gate i_t need to take values between the interval [0, 1], the activation function used to calculate these two variables is the sigmoid activation function, which is represented as σ in the formula; h_{t-1} is the output of the unit at time $t-1$ and also the input of the unit at time t. Its calculation method is as follows:

$$o_t = \sigma\left(W_o \cdot [h_{t-1},\ x_t] + b_o\right); \tag{10}$$

$$h_t = o_t \cdot \tanh\left(c_t\right). \tag{11}$$

Here, h_{t-1} refers to the output of the unit at the prior time step, and x_t is the input of the unit at the current time. These two vectors are combined as the input for computing the forget gate f_t, the input gate i_t, and the cell state update value $\widetilde{c}_t$; W_f, W_i, and W_c are the respective weights for calculating f_t, i_t, and $\widetilde{c}_t$, while b_f, b_i, and $b_c$are the biases for these three variables.

4.2. Proposed CNN-LSTM

CNN and LSTM have unique strengths in signal recognition and classification. CNN is good at focusing on the local features of signals, whereas LSTM can capture the sequential features within them. In this research, CNN and LSTM are connected in a series arrangement. First, CNN is applied to extract the local features of the signal, and then LSTM takes over to further process these extracted features.

In prediction tasks, CNN usually starts by extracting the spatial or local features from the input data. After that, the feature sequence obtained by CNN is sent to LSTM. LSTM views these feature sequences as a time series. With its gating mechanism and memory units, LSTM models the temporal relationships within the feature sequence, detecting the changing trends and patterns of features over time. Once the whole sequence is processed, LSTM produces a prediction result based on the learned features and temporal relationships. This prediction can be a single-value prediction for the next time point or a sequence prediction for multiple future time points, depending on the needs of the prediction task.

In the training phase, the prediction results are compared with the actual labels to calculate the loss function. Then, the backpropagation algorithm is used to update the parameters of both CNN and LSTM. This allows the model to constantly adjust and enhance its prediction accuracy. The prediction principle of CNN-LSTM makes full use of CNN's strong feature-extraction ability and LSTM's excellent sequence-modeling ability, giving it outstanding performance in handling data prediction tasks with both spatial and temporal features.

The prediction principle of CNN-LSTM combines the benefits of the CNN and the LSTM. The convolutional layer is the key part of CNN. It performs convolutional operations by moving the convolutional kernel across the data. Each convolutional kernel only focuses on a small area, which is called local perception. Also, the convolutional kernel uses the same weights across the entire image or data, greatly reducing the number of model parameters and the computational burden, and improving the training efficiency and generalization ability. Through the combination of multiple convolutional and pooling layers, CNN can automatically extract features at different levels from the data. The early convolutional layers can extract basic features like edges and corners. As the network layers increase, the later convolutional layers can gradually extract more advanced and abstract features, such as parts of an object and its overall shape.

The core of LSTM is its memory unit, which can store long-term information. Moreover, LSTM has three gating mechanisms: the input gate, the forget gate, and the output gate. The forget gate decides how much information to remove from the previous time step's memory unit; the input gate determines how much of the current input information to store in the memory unit; the output gate decides how much of the information in the memory unit to pass on to the current time step's hidden state. LSTM can effectively deal with long-term relationships in sequential data. For the input sequential data, it processes each time step one by one. By using the gating mechanism to selectively remember and forget information, LSTM can capture the complex relationships between different time steps in the sequence and model the long-term trends and patterns of the sequence.

The schematic of CNN-LSTM is shown in Fig. 11.

For the basic parameter settings of CNN, LSTM, and CNN-LSTM, the 'Adam' optimization algorithm was adopted, the maximum number of training epochs was set to 300, the initial learning rate was 0.001, the training set accounted for 0.8, and the test set accounted for 0.2. Table 1 shows the parameter settings of the CNN, LSTM, and CNN-LSTM models.

The evaluation index comparisons of the training set and test set for the three methods are listed in Table 2, and the main metrics are plotted in Fig. 12 as a radar chart for a very intuitive comparison.

A more intuitive comparison of the performance of the three models in terms of different evaluation metrics on the training and test sets can be made from Table 1 and Fig. 12, further validating the conclusions drawn from the previous analysis: the CNN performs well in terms of accuracy on the training

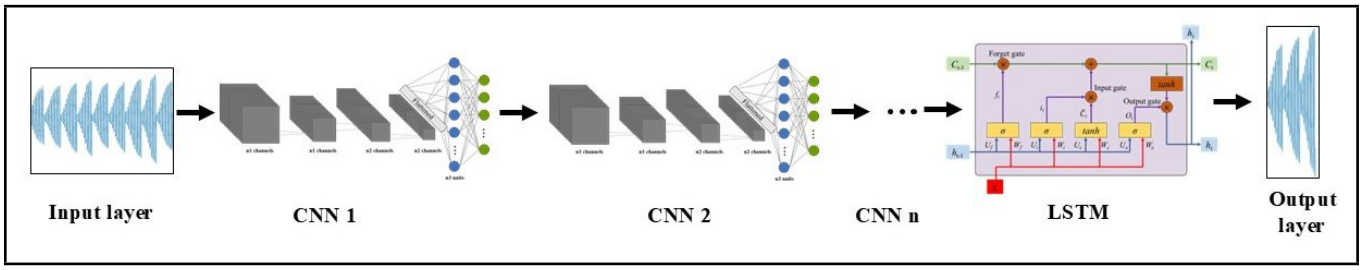

Figure 11. Schematic diagram of CNN-LSTM.

Table 1. Parameter Settings of CNN, LSTM and CNN-LSTM Models.

Model	Layer Type	Number of of Layers	Convolutional Kernel Size	Number Filters/Kernel	Pooling Size	Activation Function
CNN	Input Layer	1	-	-	-	-
	Convolutional Layer	1	[3×1]	16	-	-
	Batch Normalization Layer	1	-	-	-	-
	ReLU Layer	1	-	-	-	ReLU
	Max Pooling Layer	1	[2×1]	-	[2×1]	-
	Fully Connected Layer 1	1	-	25	-	ReLU
	Fully Connected Layer 2	1	-	1	-	-
	Regression Layer	1	-	-	-	-
LSTM	Input Layer	1	-	-	-	-
	LSTM Layer	1	-	25	-	-
	ReLU Layer	1	-	-	-	ReLU
	Fully Connected Layer	1	-	1	-	-
	Regression Layer	1	-	-	-	-
CNN-LSTM	Input Layer	1	-	-	-	-
	Convolutional Layer 1	1	[2×1]	32	-	-
	Batch Normalization Layer	1	-	-	-	-
	ReLU Layer 1	1	-	-	-	ReLU
	Max Pooling Layer 1	1	[2×1]	-	[2×1]	-
	Convolutional Layer 2	1	[2×1]	32	-	-
	Batch Normalization Layer	1	-	-	-	-
	ReLU Layer 2	1	-	-	-	ReLU
	Max Pooling Layer 2	1	[2×1]	-	[2×1]	-
	Convolutional Layer 3	1	[2×1]	32	-	-
	Batch Normalization Layer	1	-	-	-	-
	ReLU Layer 3	1	-	-	-	ReLU
	Max Pooling Layer 3	1	[2×1]	-	[2×1]	-
	Flatten Layer	1	-	-	-	-
	LSTM Layer	1	-	42	-	-
	Dropout Layer	1	-	-	-	-
	Fully Connected Layer	1	-	1	-	-
	Regression Layer	1	-	-	-	-

set, the LSTM shows relatively weaker performance across the board, and the CNN-LSTM exhibits a significant advantage in generalization ability on the test set. Without considering the training time, the comparison of the prediction results of the CNN, LSTM, and CNN-LSTM models is as follows:

4.2.1. Comparison of accuracy metrics

Training Set: The CNN has a Mean Squared Error (MSE) of 185.4936, a Root Mean Squared Error (RMSE) of 13.6196, and a Mean Absolute Error (MAE) of 8.9282. In these three metrics, the values of the CNN are smaller than those of the LSTM and CNN-LSTM, indicating that it has the smallest error margin between the predicted and true values on the training set, and thus the highest prediction accuracy. In terms of the Mean Absolute Percentage Error (MAPE), the MAPE of the CNN-LSTM is 6.3902%, lower than 7.9381% of the CNN and 10.3031% of the LSTM. This means that the CNN-LSTM has the smallest relative error on the training set.

Test Set: The CNN-LSTM has an MSE of 495.6884, an RMSE of 22.2641, and an MAE of 16.2566, all of which are lower than those of the CNN and LSTM. This demonstrates that it has higher prediction accuracy on the test set. Meanwhile, the MAPE of the CNN-LSTM is 4.6503%, much lower than that of the other two models, indicating that the relative deviation between its prediction results and the true values are the smallest.

4.2.2. Comparison of model goodness-of-fit

Training Set: The coefficient of determination (R^2) of the CNN is 99.949%, higher than 99.9311% of the LSTM and 99.9036% of the CNN-LSTM. This shows that the CNN has the best fitting effect for the training set data, and the model

Table 2. Parameter Settings of CNN, LSTM and CNN-LSTM Models.

Models	Training set					Test set				
	MSE	RMSE	MAE	MAPE	R^2	MSE	RMSE	MAE	MAPE	R^2
CNN	185.4936	13.6196	8.9282	7.9381%	0.99949	983.2891	31.3574	16.1813	8.5482%	0.998553
LSTM	255.261	15.9769	9.5646	10.3031%	0.999311	1834.25	42.8281	19.5966	16.3361%	0.997304
CNN-LSTM	376.9153	19.4143	14.4463	6.3902%	0.999036	495.6884	22.2641	16.2566	4.6503%	0.999322

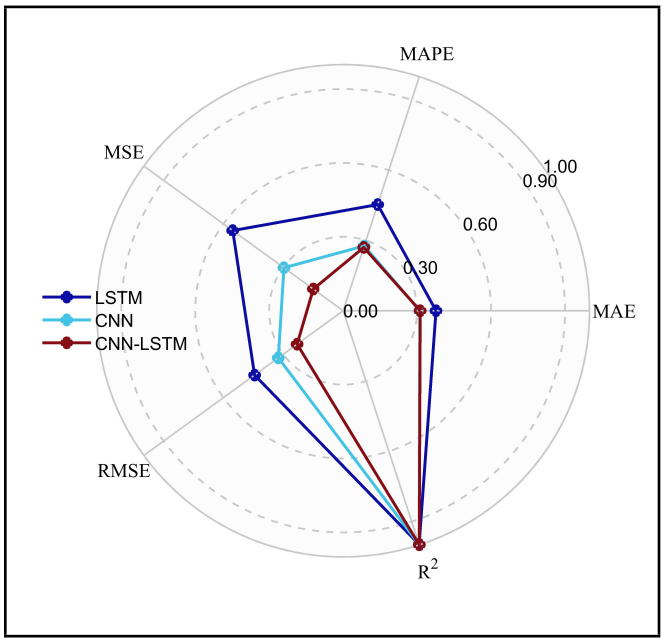

Figure 12. Radar chart of comparison of evaluation indicators for CNN, LSTM and CNN – LSTM.

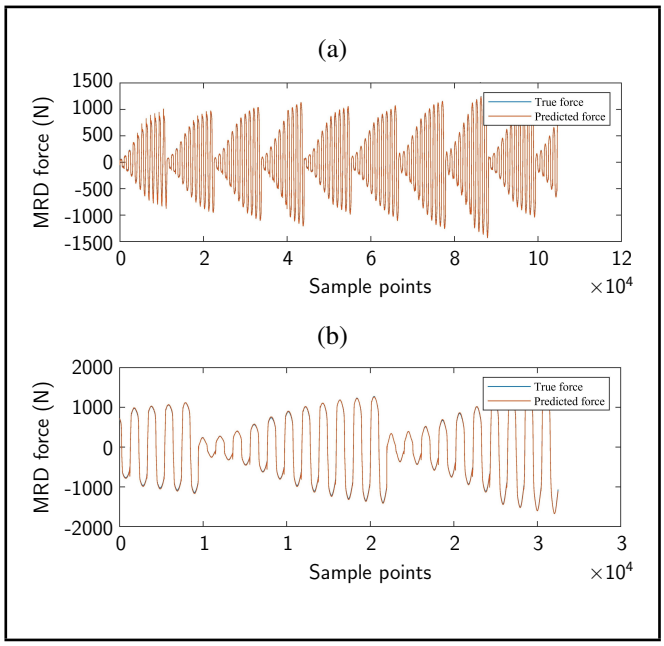

Figure 13. MRD force prediction using CNN – LSTM. (a) Training set. (b) Test set

can explain most of the variations in the training data.

Test Set: The R^2 of the CNN-LSTM is 99.9322%, higher than 99.8553% of the CNN and 99.7304% of the LSTM. This implies that the CNN-LSTM has the best fitting effect for the test set data and can better reflect the internal laws of the test data.

4.2.3. Comprehensive comparison

The CNN shows outstanding prediction accuracy on the training set. Most of its error metrics (except for MAPE) indicate a high prediction accuracy for the training data, and it also has the highest goodness-of-fit. This suggests that it can effectively learn the characteristics of the training data.

The LSTM shows relatively poor performance in all indicators on both the training and test sets. Whether it is prediction accuracy or goodness-of-fit, it is inferior to the CNN and CNN-LSTM.

The CNN-LSTM demonstrates strong generalization ability on the test set. Its error metrics perform best on the test set, and its goodness-of-fit is also the highest. Although its accuracy on the training set is not the optimal, it can better adapt to unknown data, which indicates that it may have better performance in practical applications.

Figure 13 shows the comparison between the true MRD force and the predicted one on the training set and test set based on CNN-LSTM.

5. CONCLUSION

This study successfully applies deep learning technology to the identification and prediction of MRD output force, effectively overcoming the limitations of traditional analysis methods and opening new avenues for research and applications in this field.

In terms of methodological innovation, the CWT was used to preprocess the time series of MRD output. This transformed the data into two - dimensional time-frequency diagrams rich in time-frequency information, providing more valuable input for the CNN. As a result, the CNN was able to extract the key features of MRD force more precisely, enabling efficient identification of MRD output. Meanwhile, the constructed CNN-LSTM model fully integrated the powerful local feature - extraction ability of the CNN and the excellent time-series modeling advantage of the LSTM, offering a more reliable method for predicting MRD force.

In the performance evaluation of the MRD output prediction models, the CNN-LSTM demonstrated unique advantages. Although the CNN had high accuracy on the training set, with excellent performance in various error metrics and the highest goodness-of-fit, which allowed it to learn the characteristics of the training data well, this also posed a risk of overfitting to the training data. The LSTM, on the other hand, had relatively weak overall performance on both the training and test sets, with certain limitations in capturing local signal features and handling complex data. In contrast, the CNN-LSTM model exhibited strong generalization ability on the test set. Its error metrics were the best among the models on the test set, and its

goodness-of-fit was also the highest. This fully demonstrated its ability to effectively integrate the local feature-extraction ability of the CNN and its own excellent time-series modeling ability. Even when faced with unknown data that differed significantly from the training data, the CNN-LSTM could accurately capture data features and predict MRD output precisely, showing remarkable adaptability. This advantage gives the CNN-LSTM great potential in practical engineering applications. In fields such as building structure seismic protection, bridge vibration reduction, and vehicle suspension systems, the CNN-LSTM is expected to provide more reliable support for optimizing MRD control strategies through its powerful performance, thereby effectively improving vibration control efficiency and advancing technological development in related fields.

The results of this study have significant guiding implications for engineering practice. They provide a scientific basis for optimizing MRD control strategies, contribute to significantly enhancing vibration control effectiveness, and promote the in-depth application and innovative development of deep learning technology in the field of MRD vibration control.

Looking ahead, there is still much work to be done. On one hand, the model structure and parameters can be further optimized, and the model performance under more complex working conditions and environments can be explored to enhance the model's adaptability and robustness. On the other hand, attempts can be made to extend the methods of this study to more types of intelligent material devices or systems, expanding the application scope of the research results. Additionally, by combining with other cutting-edge technologies such as reinforcement learning and transfer learning, it is expected to further improve the performance and application value of the model, providing more powerful technical support for the development of related fields.

Looking ahead, the MRD output force identification and prediction technology proposed in this research holds promising applications across multiple critical domains. In the realm of smart buildings and seismic protection, the integration of CWT for time-frequency analysis with the real-time predictive capabilities of the CNN-LSTM model will enable adaptive regulation of dynamic damping forces in high-rise structures, bridges, and similar infrastructures during earthquakes or strong winds. By precisely identifying vibration patterns and pre-emptively predicting force outputs, this technology will significantly enhance structural safety against seismic and wind-induced forces. Within the automotive and rail transit sector, the technology will adeptly respond to dynamic scenarios such as varying road undulations and track irregularities, enabling rapid optimization of MRD damping forces in suspension systems. Leveraging the 100% accuracy of CWT-CNN in vibration pattern recognition, it will facilitate real-time monitoring of vehicle operational status and early fault detection, thereby improving ride comfort and operational stability. In precision manufacturing and research facilities, the model's high sensitivity in identifying and predicting micro-vibrations will deliver ultra-precise vibration control for vibration-critical environments like semiconductor lithography equipment and quantum laboratories. This will effectively suppress environmental micro-disturbances, ensuring the precision of experiments and manufacturing processes. Furthermore, when integrated with IoT technologies, the technology will become an integral part of smart city vibration monitoring networks. It will enable comprehensive vibration status sensing and coordinated control across building complexes and transportation hubs, providing data-driven decision support for the health management and disaster prevention of large-scale infrastructure. These application scenarios all capitalize on the technology's core strengths in processing complex dynamic signals, achieving high-precision predictions, and demonstrating strong generalization capabilities. They are poised to drive the industrial upgrading of MRDs in the field of intelligent vibration control.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDINGS

The authors also wish to express gratitude to the National Key Research and Development Program (Grant no. 2024YFF0508104), *Research and Application of Vibration and Structural Safety Control Methods for Complicated Underground Rail Transit.* The authors also wish to acknowledge the Key Youth Fund of SINOMACH (Grant no. QNJJZD – 2022 - 04), *Research and application of key technologies for micro-nano environmental vibration control of Chinese major science and technology infrastructures*, and the major technology development project of SINOMACH (Grant no. ZDZX2025-07), *Research on key technologies for industrial upstairs engineering design*, *The research fund of the Academician Workstation of China Academy of Building Research Co., Ltd., Research on the Mechanism of Engineering Vibration and Seismic Vibration Response and Key Performance Improvement Technologies for Building Structures.*

REFERENCES

[1] Bhowmik, K. and Debnath, N. Semi-active vibration control of soft-storey building with magnetorheological damper under seismic excitation, *Journal of Vibration Engineering & Technologies*, **12** (4), 6943–6961, (2024). https://doi.org/10.1007/s42417-024-01292-5

[2] Tell, S., Andersson, A., Najafi, A., Spencer Jr, B. F. and Karoumi, R. Real-time hybrid testing for efficiency assessment of magnetorheological dampers to mitigate train-induced vibrations in bridges, *International Journal of Rail Transportation*, **10** (4), 436–455, (2022). https://doi.org/10.1080/23248378.2021.1954560

[3] Josee, M., Sosthene, K. and Turabimana, P. Review of semi-active suspension based on Magneto-rheological

damper, *Engineering Perspective*, **1** (2), 38–51, (2021). https://doi.org/10.29228/eng.pers.50853

[4] Zhang Y, Guo J, Yang J, et al. Recent structural developments and applications of magnetorheological dampers (MRD): a review, *Magnetochemistry*, **9** (4), 90, (2023). https://doi.org/10.3390/magnetochemistry9040090

[5] Javanbakht, M., Cheng, S. and Ghrib, F. Semi-active adaptive control of stay cable vibrations using MR dampers, *Structural Control and Health Monitoring*, **29** (12), e3121, (2022). https://doi.org/10.1002/stc.3121

[6] Liao, G., Gong, X. and Xuan, S. Phase based stiffness tuning algorithm for a magnetorheological elastomer dynamic vibration absorber, *Smart Materials and Structures*, **23** (1), 015016, (2014). https://doi.org/10.1088/0964-1726/23/1/015016

[7] Ren, H., Zhao, D., Dong, L., Liu, S., Yang, J., Zhao, T. and Fan, Y. Deep learning accelerates reverse design of Magnetorheological elastomer, *Composites Science and Technology*, **265**, 111148, (2025). https://doi.org/10.1016/j.compscitech.2025.111148

[8] Ren, H., Zhao, D., Dong, L., Liu, S. and Yang, J. A physics-guided deep learning model for predicting the magneto-induced mechanical properties of magnetorheological elastomer: Small experimental data-driven, *Composites Science and Technology*, **253**, 110653, (2024). https://doi.org/10.1016/j.compscitech.2024.110653

[9] Guo, Q., Yang, X., Li, K. and Li, D. Parameters identification of magnetorheological damper based on particle swarm optimization algorithm, *Engineering Applications of Artificial Intelligence*, **143**, 110016, (2025). https://doi.org/10.1016/j.engappai.2025.110016

[10] Bedoya-Zambrano, D., Lara-Valencia, L. and Blandón-Valencia, J. Optimization of control forces in a three-dimensional frame with magnetorheological dampers using a hybrid algorithm, *Journal of Building Engineering*, **100**, 111697, (2025). https://doi.org/10.1016/j.jobe.2024.111697

[11] Lv, J., Wu, M., Zhao, T., He, J. and Wei, Y. Accurate prediction of magnetorheological damper characteristics based on a new rheological constitutive model, *Structures*, **50**, 108–117, (2023). https://doi.org/10.1016/j.istruc.2023.02.025

[12] Yoo, Y. and Jeong, S. Vibration analysis process based on spectrogram using gradient class activation map with selection process of CNN model and feature layer, *Displays*, **73**, 102233, (2022). https://doi.org/10.1016/j.displa.2022.102233

[13] Chen, H. Y. and Lee, C. H. Deep Learning Approach for Vibration Signals Applications, *Sensors*, **21** (11), 3929, (2021). https://doi.org/10.3390/s21113929

[14] Sony, S., Gamage, S., Sadhu, A. and Samara Bandu, J. Vibration-based Multiclass Damage Detection and Localization using Long Short-Term Memory Networks, *Structures*, **35**, 436–451, (2022). https://doi.org/10.1016/j.istruc.2021.10.088

[15] Vos, K., Peng, Z., Jenkins, C., Shahriar, M. R., Borghesani, P. and Wang, W. Vibration-based Anomaly Detection using LSTM/SVM Approaches, *Mechanical Systems and Signal Processing*, **169**, 108752, (2022). https://doi.org/10.1016/j.ymssp.2021.108752

[16] Eang, C. and Lee, S. Predictive maintenance and fault detection for motor drive control systems in industrial robots using CNN-RNN-based observers, *Sensors*, **25** (1), 25, (2024). https://doi.org/10.3390/s25010025

[17] Lee, J.-H. and Hong, J.-K. Comparative Performance Analysis of RNN Techniques for Predicting Concatenated Normal and Abnormal Vibrations, *Electronics*, **12** (23), 4778, (2023). https://doi.org/10.3390/electronics12234778

[18] Łuczak, D. Machine fault diagnosis through vibration analysis: continuous wavelet transform with complex morlet wavelet and time–frequency rgb image recognition via convolutional neural network, *Electronics* (2079-9292), **13** (2), (2024). https://doi.org/10.3390/electronics13020452

[19] Zhao, H., Liu, J., Chen, H. et al. Intelligent diagnosis using continuous wavelet transform and gauss convolutional deep belief network, *IEEE Transactions on Reliability*, **72** (2), 692–702, (2022). https://doi.org/10.1109/TR.2022.3180273

[20] Zafar, A., Aamir, M., Mohd Nawi, N. et al. A comparison of pooling methods for convolutional neural networks, *Applied Sciences*, **12** (17), 8643, (2022). https://doi.org/10.3390/app12178643

[21] Hyun J, Seong H, Kim E. Universal pooling–a new pooling method for convolutional neural networks, *Expert Systems with Applications*, **180**, 115084, (2021). https://doi.org/10.1016/j.eswa.2021.115084

[22] Cai, T. T. and Ma, R. Theoretical foundations of t-sne for visualizing high-dimensional clustered data, *Journal of Machine Learning Research*, **23** (301), 1–54, (2022). https://doi.org/10.48550/arXiv.2105.07536

[23] Agis, D. and Pozo, F. Vibration-based Structural Health Monitoring Using Piezoelectric Transducers and Parametric t-SNE, *Sensors* (Basel), **20** (6), 1716, (2020). https://doi.org/10.3390/s20061716

[24] Kong, X., Li, X., Zhou, Q. et al. Attention recurrent autoencoder hybrid model for early fault diagnosis of rotating machinery, *IEEE Transactions on Instrumentation and Measurement*, **70**, 1–10, (2021). https://doi.org/10.1109/TIM.2021.3051948