

Article

Prediction of Residual Electrical Life in Railway Relays Based on Convolutional Neural Network Bidirectional Long Short-Term Memory

Shuxin Liu, Yankai Li *, Shuyu Gao, Chaojian Xing, Jing Li and Yundong Cao

School of Electrical Engineering, Shenyang University of Technology, Shenyang 110870, China; liushuxin@sut.edu.cn (S.L.); gaoshuyu@smail.sut.edu.cn (S.G.); 15524325220@163.com (C.X.); lijing@sut.edu.cn (J.L.); caoyd@sut.edu.cn (Y.C.)

* Correspondence: 15524325220@139.com

Abstract: In this paper, we address several issues with existing methods for predicting the residual electrical life of railway relays. These issues include the difficulty of single feature prediction in characterizing the degradation process, the neglect of temporal and backward–forward correlations in the degradation process, and low prediction accuracy. To overcome these challenges, we propose a novel approach that combines convolutional neural networks (CNNs) and bidirectional long short-term memory (BiLSTM) to facilitate the life prediction of railway relays and provide an accurate data basis for the maintenance of railway relays. Firstly, we collected voltage and current signals from railway relay electrical life tests and extracted feature parameters that captured the relay’s operating state. Next, we applied Spearman correlation coefficient analysis combined with random forest importance analysis to perform double-feature selection. This process eliminates redundant feature parameters and identifies the optimal feature subset. Finally, we constructed a convolutional neural network bidirectional long short-term memory (CNN BiLSTM) prediction model to accurately predict the remaining electrical life of the railway relay. Through our analysis of the prediction results, we observed that the CNN BiLSTM model achieves an effective prediction accuracy of 96.3%. This accuracy is significantly higher, more stable, and more practical compared to other prediction models such as recurrent neural networks (RNNs), long short-term memory (LSTM), and BiLSTM models. Overall, our proposed CNN BiLSTM model offers higher accuracy, better stability, and greater practicality in predicting the remaining electrical life of railway relays.

Keywords: railway relay; random forest; Spearman correlation coefficient analysis; CNN; BiLSTM; life prediction

Citation: Liu, S.; Li, Y.; Gao, S.; Xing, C.; Li, J.; Cao, Y. Prediction of Residual Electrical Life in Railway Relays Based on Convolutional Neural Network Bidirectional Long Short-Term Memory. *Energies* **2023**, *16*, 6357. <https://doi.org/10.3390/en16176357>

Academic Editor: Daniel-Ioan Stroe

Received: 6 August 2023

Revised: 25 August 2023

Accepted: 29 August 2023

Published: 1 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Railway relays are essential switching devices in railway equipment, serving as specialized multi-way switches for controlling circuits by opening or closing them [1]. They play a crucial role in signal transmission, logic control, and other functions in trains. The stability and safety of relays are vital for ensuring the normal operation of railway signal systems [2]. Predicting the remaining electrical life of railway relays allows us to identify critical time points for performance testing, thus facilitating maintenance, improving economic efficiency, reducing failure rates, enhancing safety, and ensuring the safe and stable operation of the entire train [3,4].

As the performance requirements for relays have increased, research on the prediction of relay remaining electrical life has made progress. Reference [5] proposed an electromagnetic relay life prediction method based on a weighted optimization combination model, using degradation parameters of electromagnetic relays for the prediction process.

Reference [6] utilized rough set theory to extract a set of lifespan-related feature parameters from various parameters and developed a relay lifespan decision making method, completing the prediction process based on initial product life state information of the same model. Reference [7] addressed the issue of non-stationary time series for relay performance parameters and achieved lifespan prediction through wavelet packet transform and RBF neural networks. Reference [8] incorporated experimentally obtained relay overrun time data into a partitioned regression model to predict relay lifespan. Reference [9] proposed a support vector machine-based prediction method for electromagnetic relay storage life. Reference [10] systematically analyzed the common failure modes and their causes of electromagnetic relays during the working process, thus providing a method for life determination. Reference [11] evaluated the relay life by monitoring the working status of relay contacts. Reference [12] proposed a one-dimensional convolutional neural network-based automotive electromagnetic relay life prediction method with a dual self-attention mechanism. Reference [13] estimated the relay life at room temperature by using unknown variables in the Arrhenius model. Reference [14] implements the life prediction process by predicting the fracture cycle. Reference [15] proposed an electromagnetic relay life prediction method based on improved DBN and Softmax regression. Reference [16] established a balanced force-based electromagnetic relay life prediction model based on a failure mechanism.

Summarizing existing research results, three aspects need further improvement in the prediction of railway relay remaining electrical life. Firstly, existing prediction methods mostly rely on single-variable prediction, which fails to adequately represent the relay's performance degradation process, leading to significant prediction errors. Secondly, the neglect of the relay's performance degradation process as a long time series problem hinders the reasonable utilization of its temporal correlations during the lifespan prediction process, affecting the accuracy of predictions. Thirdly, most existing prediction methods are based on traditional prediction algorithms, such as regression prediction, which is widely used in the references. However, its prediction accuracy is low, and the excessive amount of data can seriously affect the prediction accuracy, making it difficult to meet the requirements of practical applications.

Recently, deep learning-based prediction methods have made progress in various research areas such as power load forecasting [17], mechanical lifespan prediction [18], and battery lifespan prediction [19]. CNN, compared to other mathematical models, has stronger representation learning capabilities as it can learn through layered structures from input feature parameter data, with more noticeable feature extraction effects [20]. LSTM, an extension of RNN, introduced forget, input, and output gates, along with a cell state, greatly alleviating the vanishing or exploding gradient problem when dealing with time series with long-term dependencies [21]. BiLSTM, composed of forward and backward LSTMs, can simultaneously capture information from both directions, providing a clear advantage in capturing bidirectional temporal sequences.

Addressing the shortcomings of existing research and combining the advantages of CNN and BiLSTM models, this paper proposes a CNN BiLSTM-based railway relay residual electric life prediction method, which will realize the accurate life prediction of railway relays through the optimal feature parameter subset obtained by feature filtering, providing an accurate data basis for the maintenance of the railway relay and improving the safety and reliability of the railway system. The method will provide an accurate data basis for the maintenance of railway relays and improve the safety and reliability of the railway system. This method can directly perform life prediction using the feature parameter data obtained from the full life cycle test of railway relays, without the need to build accurate physical or statistical models [22]. Therefore, the deep learning life prediction method is more advantageous in solving the long time series prediction problem of relay remaining life.

This study first designed a railway relay electric life test platform to complete the test process, collect voltage and current signals, and extract feature parameters. Then, through

Spearman correlation coefficient analysis combined with random forest importance analysis, the feature parameter selection process was completed and the optimal feature parameter subset was obtained. Finally, the railway relay life prediction process was completed through CNN BiLSTM, and multiple methods were compared to prove its feasibility. The overall prediction process is based on multiple feature parameters, which can better characterize the degradation process of railway relays. The combined model reduces computational complexity; it can consider the correlation before and after the degradation process of railway relays, capture bidirectional time series information, and improve prediction accuracy and stability, thus achieving accurate life prediction of railway relays.

2. Basic Theoretical Foundation

2.1. Spearman Correlation Coefficient Analysis

The Spearman correlation coefficient reduces the influence of outliers during the process of correlation analysis, and it also exhibits stronger robustness in capturing non-linear relationships [23,24]. The formula for calculating the Spearman correlation coefficient is:

$$\rho_{xy} = \frac{\sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_{k=1}^n (x_k - \bar{x})^2 \sum_{k=1}^n (y_k - \bar{y})^2}} \quad (1)$$

In the equation, x_k represents the rank of the k -th element in the vector X , and $\bar{x}$ denotes the average rank of all elements in vector X . Similarly, y_k represents the rank of the k -th element in the vector Y , and $\bar{y}$ denotes the average rank of all elements in vector Y . The Spearman correlation coefficient's values correspond to the relationship exhibited between vectors x and y , as shown in Table 1.

Table 1. Spearman's correlation coefficient.

ρ_{xy} Value Range	the Correlation between x and y
$0 < \rho_{xy} \leq 0.2$	No correlation
$0.2 < \rho_{xy} \leq 0.4$	Weak correlation
$0.4 < \rho_{xy} \leq 0.7$	Moderate correlation
$0.7 < \rho_{xy} \leq 1$	Strong correlation

2.2. Importance Analysis of the Random Forest Features

Random forest is a parallel ensemble learning model that uses decision trees as base classifiers [25]. In this study, the feature importance analysis is conducted using the random forest's mean decrease accuracy method, with the evaluation metric being the out-of-bag (OOB) error rate. The importance VIM_j^{OOB} of a variable X_j in the i -th tree is denoted as:

$$VIM_j^{(OOB)} = \frac{\sum_{p=1}^{n_0^i} I(Y_p = Y_p^i)}{n_0^i} - \frac{\sum_{p=1}^{n_0^i} I(Y_p = Y_{p,\pi_j}^i)}{n_0^i} \quad (2)$$

In the equation, n_0^i represents the number of observations in the i -th tree's OOB data, $Y_p \in \{0,1\}$ is the true outcome of the p -th observation, $Y_p^i \in \{0,1\}$ represents the predicted outcome of the i -th tree for the p -th observation in the OOB data before random permutation, and $Y_{p,\pi_j}^i \in \{0,1\}$ represents the predicted outcome of the i -th tree for the p -

th observation in the *OOB* data after random permutation. I denotes the indicator function, which takes the value of 1 when the two values are equal and 0 otherwise. When a variable j does not appear in the i -th tree, $VIM_j^{(OOB)} = 0$.

The replacement importance of a variable X_j in RF is defined as:

$$VIM_j^{(OOB)} = \frac{\sum_{i=1}^n VIM_{ij}^{(OOB)}}{n} \quad (3)$$

In the equation, n represents the number of classification trees in the random forest (RF).

When performing feature importance analysis through random forest, the total importance sum is equal to 1. When the accumulated importance of feature parameters reaches 0.8 or above, it is considered that these feature parameters have representative feature information and have a significant impact on the model's predictions. Therefore, they can be used as the basis for data in the predictive model.

2.3. Convolutional Neural Network

CNN has excellent feature extraction capabilities, and it consists of an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer [26–28]. The schematic diagram of the convolutional neural network model is shown in Figure 1.

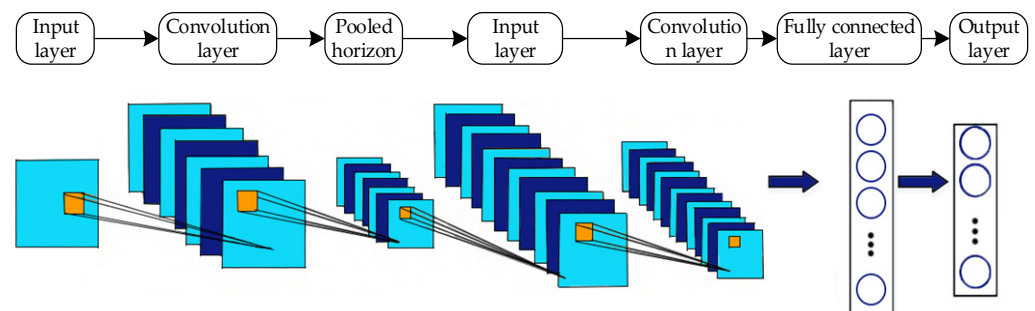


Figure 1. A schematic diagram of the convolutional neural network.

The convolutional layer performs convolution operations on the input features using convolutional kernels, which can be represented as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \quad (4)$$

In the equation, x_j^l represents the j -th feature map output after the l -th convolutional operation, f is the chosen activation function for the model, k_{ij}^l is the input for the l -th layer, b_j^l is the bias parameter for the l -th layer, and M_j is the convolutional kernel for the input feature vector.

The role of the activation function is to construct the output feature vector. *ReLU* function is chosen as the activation function for this model to prevent gradient vanishing. It can be represented as follows:

$$R_{eLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (5)$$

In the equation, x represents the output data of the convolution operation procedure.

The pooling layer can reduce the size of the model and improve the computational speed of the predictive model. The pooling operation chosen for this model is max pooling, which can be represented as follows:

$$P_i^{l+1}(j) = \max_{(j-1)W+1 \leq t \leq jW} \{q_i^l(t)\} \quad (6)$$

In the equation, $q_i^l(t)$ represents the value of the t -th neuron of the i -th feature vector in the l layer, $t \in [(j-1)W+1, jW]$, $P_i^{l+1}(j)$ represents the values corresponding to the neurons in the $l+1$ -th layer, and W represents the width of the pooling field.

The fully connected layer will map the learned features into the label space, and it can be represented as follows:

$$f(x) = W \times x + b \quad (7)$$

In the equation, W represents the weight matrix of the fully connected layer, X represents the input, and b represents the bias vector.

2.4. Bidirectional Long Short-Term Memory Network

BiLSTM is composed of two layers of bidirectional LSTM, which enables it to consider both forward and backward context information. It has been widely applied in language modeling, sentiment analysis, machine translation, sequence labeling, time series prediction, and other areas.

LSTM networks address the issues of vanishing and exploding gradients that may occur during training in traditional RNNs. They enhance the capability of handling long-term time series data [29,30]. The recurrent structure and internal architecture of LSTM networks are illustrated in Figure 2.

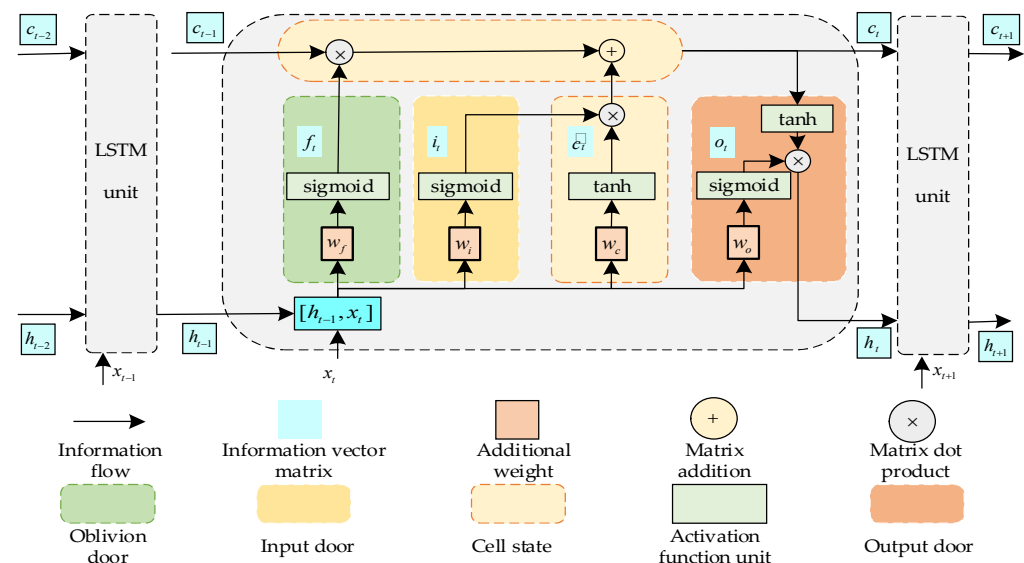


Figure 2. The LSTM loop structure and the internal structure.

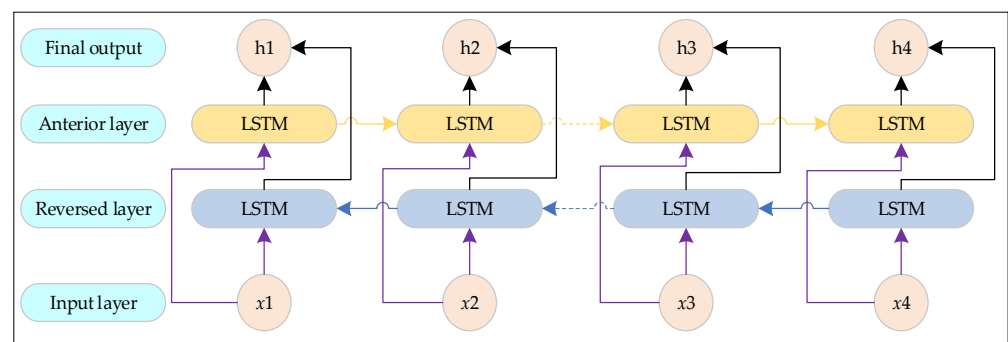
The mathematical calculation formula of LSTM is shown below (Table 2):

Table 2. LSTM computational formula.

LSTM Internal Structure	Expression Formula
Oblivion door value f_t	$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$
Input door value i_t	$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$
Output value o_t	$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$
The cell state at the current time c_t	$c_t = f_t \circ c_{t-1} + i_t \circ \theta_t$
The current input cell status $\theta_t \in [0,1]$	$\theta_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$
Final output o_t	$h_t = o_t \circ \tanh(c_t)$

In the above table, σ is the sigmoid activation function, W_f is the forgotten door weight matrix, $[h_{t-1}, x_t]$ is a combined matrix of the upper moment output h_{t-1} and the current moment input x_t , b_f is the forgotten door bias item, W_i is the input door weight matrix, b_i is the input gate bias, $\tanh$ is hyperbolic tangential activation function, W_c is the input unit state weight matrix, b_c is the bias of unit state, $\circ$ is the multiplication of the matrix elements, c_{t-1} is the last moment unit state, W_o is the output door weight matrix, b_o is the output door bias term, and h_t is the final output value.

BiLSTM incorporates an additional backward LSTM layer into the LSTM structure, allowing it to simultaneously consider both forward and backward context information. By capturing contextual information associated with the target event from both directions, BiLSTM can improve the model's performance [31,32]. The recurrent structure of BiLSTM is illustrated in Figure 3.

**Figure 3.** Circular structure of BiLSTM.

2.5. Parameter Setting and Evaluation Index

(1) Parameter Setting

The remaining action data of railway relays are taken as the label of the prediction model, and the optimal feature subset obtained through double feature selection is taken as the input of the model. The first 70% of the data are selected as the training set, 70% to 85% of the data are taken as the validation set, and the last 15% of the data are taken as the test set. The parameters of the convolutional part are set as follows: the number of convolutional layers is 1, the number of kernels is 6, and the activation function is ReLU. The parameters of the BiLSTM part are set as follows: the loss function is selected as the mean square error (MSE) function, the time step is 48, there are two BiLSTM layers, the number of cells is 32 and 8, the iteration times are 40, the optimizer is Adam, and the learning rate is 0.01.

(2) Evaluation Metrics

The following metrics are used as evaluation indicators to assess the performance of the model: root mean square error (*RMSE*), mean absolute error (*MAE*), coefficient of determination (R^2), maximum error rate, effective accuracy, and computational time. The formulas for calculating *RMSE*, *MAE*, and R^2 are as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (9)$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

where n is the number of samples of the data set, y_i is the actual value of the remaining break times, $\bar{y}$ is the average of the actual value, and $\hat{y}_i$ is the final predicted output value.

RMSE (root mean square error) and *MAE* (mean absolute error) are inversely related to the prediction accuracy, meaning that lower *RMSE* and *MAE* values indicate better prediction performance. On the other hand, R^2 is directly related to the prediction effectiveness. A higher R^2 value indicates a better fit of the model to the actual data [33,34].

3. Test Platform Design and Parameter Extraction

3.1. Design and Construction of the Test Platform

In this experiment, the 3RH211-2KF40 (Siemens, Berlin, Germany) model railway relay was selected as the test object. The experiment utilized the combination of communication serial ports to connect the NI data acquisition card with LabVIEW virtual instrument technology. The data acquisition was carried out using the X-Series USB6356 synchronous data acquisition card from National Instruments (Austin, TX, USA) and voltage and current sensors from LEM (Geneva, Switzerland). This setup facilitated the collection of signals related to the main contact voltage, main contact current, and coil voltage and current during the entire make-and-break process of the railway relay, thus accomplishing the design and construction of the railway relay's electrical life test platform. The design conforms to the requirements of the GB14048.5-2020 national standard [35]. Specific experimental parameters are listed in Table 3. The schematic diagram of the experimental platform is illustrated in Figure 4.

Table 3. Experimental conditions.

Test Condition	Specific Parameter
Coil voltage	DC 110 V
Load voltage	DC 24 V
On condition	$I/I_e = 1, U/U_e = 1, T_{0.95} = 1ms$
Load current	10 A
Sampling frequency	200 K/Hz
Operating frequency	360 times/h
Breaking condition	$I_c/I_e = 1, U/U_e = 1, T_{0.95} = 1ms$
Operating voltage	DC 12 V
Load type	Resistive load

In the above table, I represents the switching current, I_e represents the rated operating current, U represents the externally applied voltage, U_e represents the rated operating voltage, I_c represents the breaking current, and $T_{0.95}$ represents the time taken to reach 95% of the steady-state current.

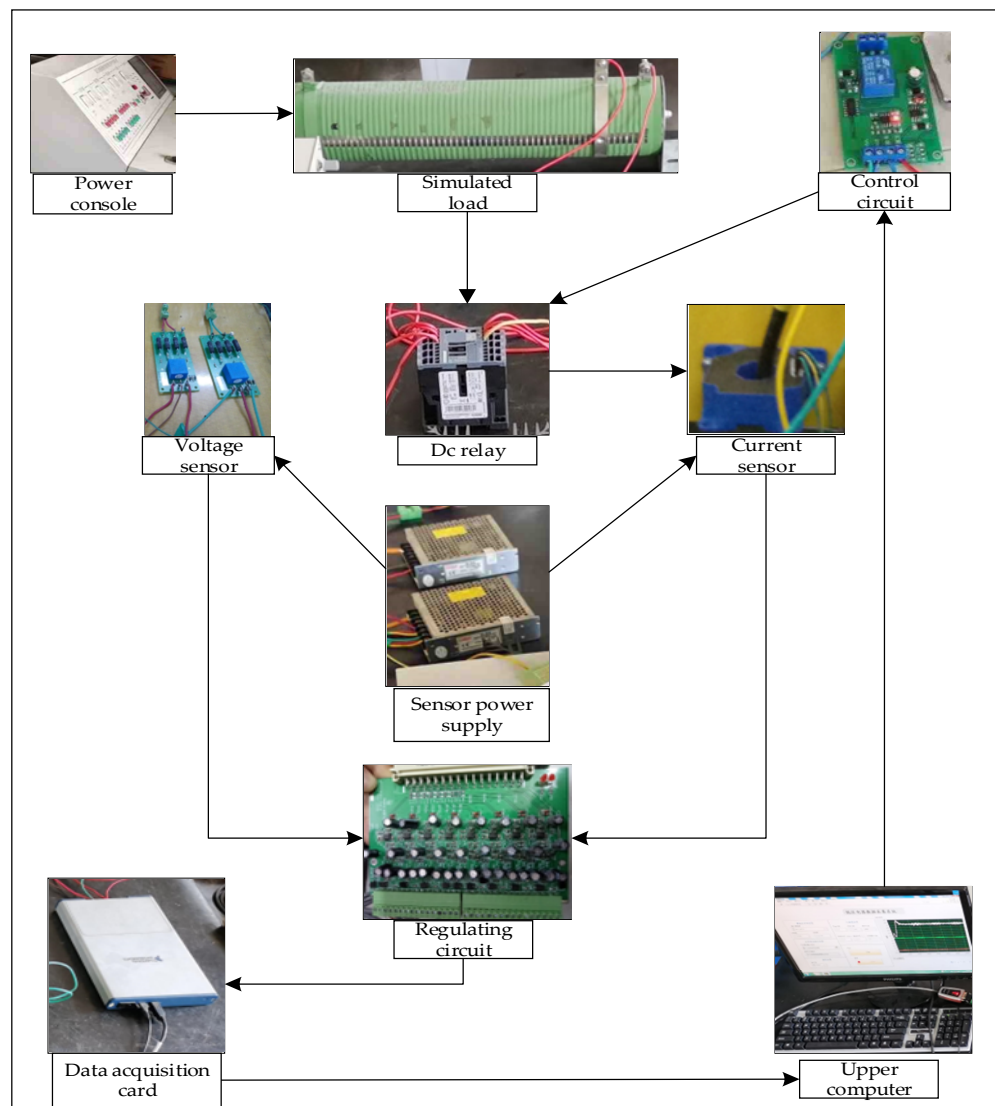


Figure 4. A schematic diagram of the test platform.

3.2. Feature Parameter Extraction

Feature parameters serve as the basis for a railway relay's electrical life prediction, and using multiple feature parameters allows for a better representation of the relay's degradation process. The extraction of these feature parameters is based on the voltage and current signals during the make-and-break process of the railway relay. The waveform variations are depicted in Figures 5 and 6.

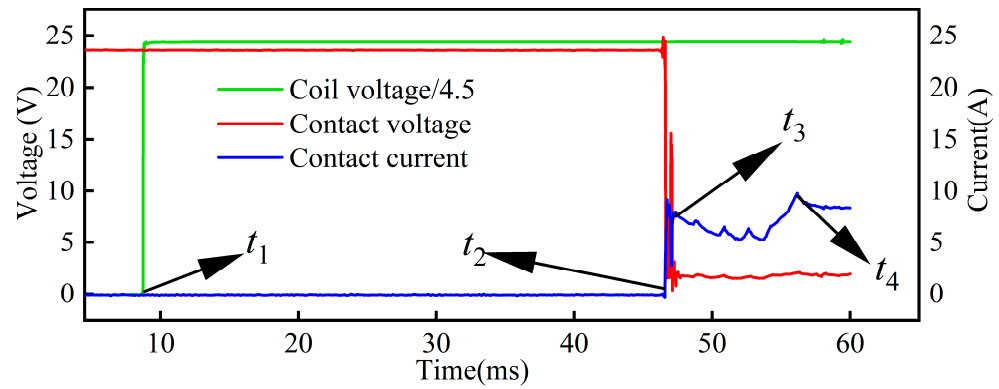


Figure 5. Waveform diagram of relay suction.

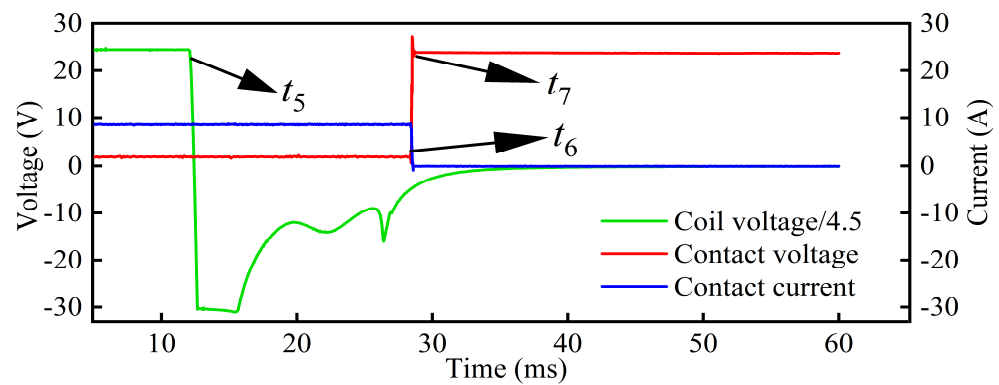


Figure 6. Waveform diagram of relay breaking.

Compared to single feature parameter prediction, electrical life prediction based on multiple feature parameters can better capture the interrelationships between features and exhibits stronger sequence modeling capabilities. Therefore, this study extracted eight types of feature parameters. The extraction process is as follows:

(1) Contact resistance extraction:

$$R_n = \frac{\sqrt{\frac{1}{N} \sum_{n=1}^N u_n^2}}{\sqrt{\frac{1}{N} \sum_{n=1}^N i_n^2}} = \sqrt{\frac{\sum_{n=1}^N u_n^2}{\sum_{n=1}^N i_n^2}} \quad (11)$$

where N represents the number of data points collected when the contacts are stably closed and conducting. u_n represents the voltage across the contacts, and i_n represents the current flowing through the contacts.

(2) Suction time extraction:

$$t_x = t_2 - t_1 \quad (12)$$

where t_1 is the moment when the coil is energized, and t_2 is the first contact time between the dynamic and static contacts.

(3) Bounce time extraction:

$$t_i = t_3 - t_2 \quad (13)$$

where t_3 is the moment when significant voltage and current fluctuations end.

(4) Overshoot time extraction:

$$t_c = t_4 - t_2 \quad (14)$$

where t_4 is the moment when stable contact closure is achieved.

(5) Release time extraction:

$$t_s = t_6 - t_5 \quad (15)$$

where t_5 represents the moment when the coil power is disconnected, and t_6 represents the moment when the arcing starts.

(6) Arc energy extraction:

$$E = \sum_{n=t_6}^{t_7} u_n \cdot i_n \cdot \Delta t = \frac{\sum_{n=t_6}^{t_7} u_n \cdot i_n}{f_s} \quad (16)$$

where t_7 is the arc extinguishing time, Δt is the time interval between adjacent sampling points, and f_s is the sampling rate.

(7) Arc time extraction:

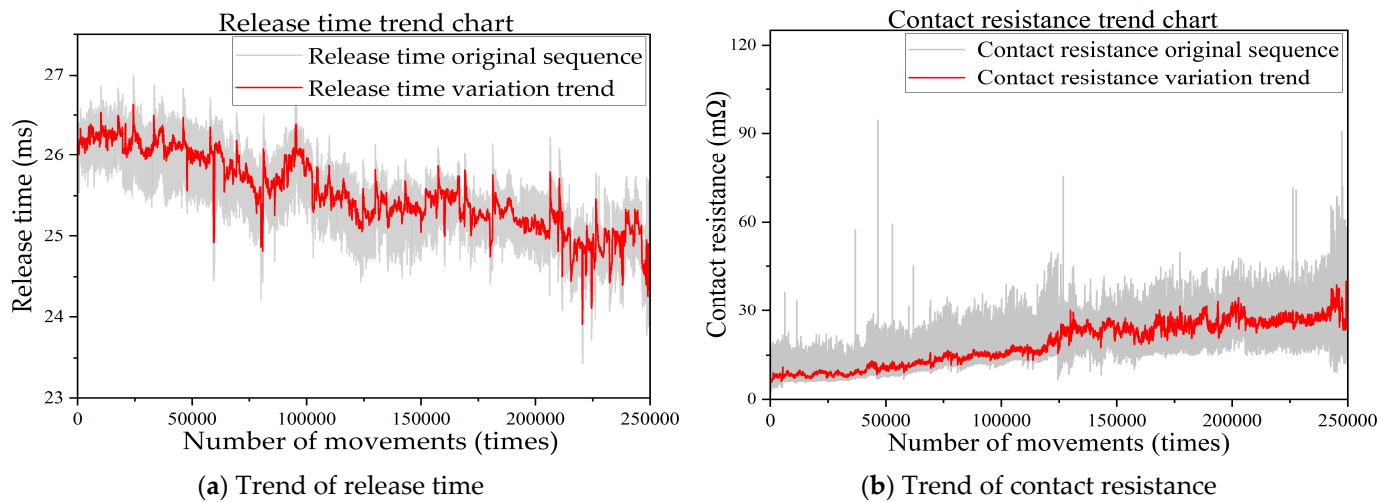
$$t_{arc} = t_7 - t_6 \quad (17)$$

(8) Release voltage extraction:

$$u_f = \frac{1}{N} \sum_{n=t_6}^N u_{nx} \quad (18)$$

where u_{nx} represents the coil voltage at different numbers of collected points when the contacts are opened.

The extracted feature parameters and their variations are shown in Figure 7.



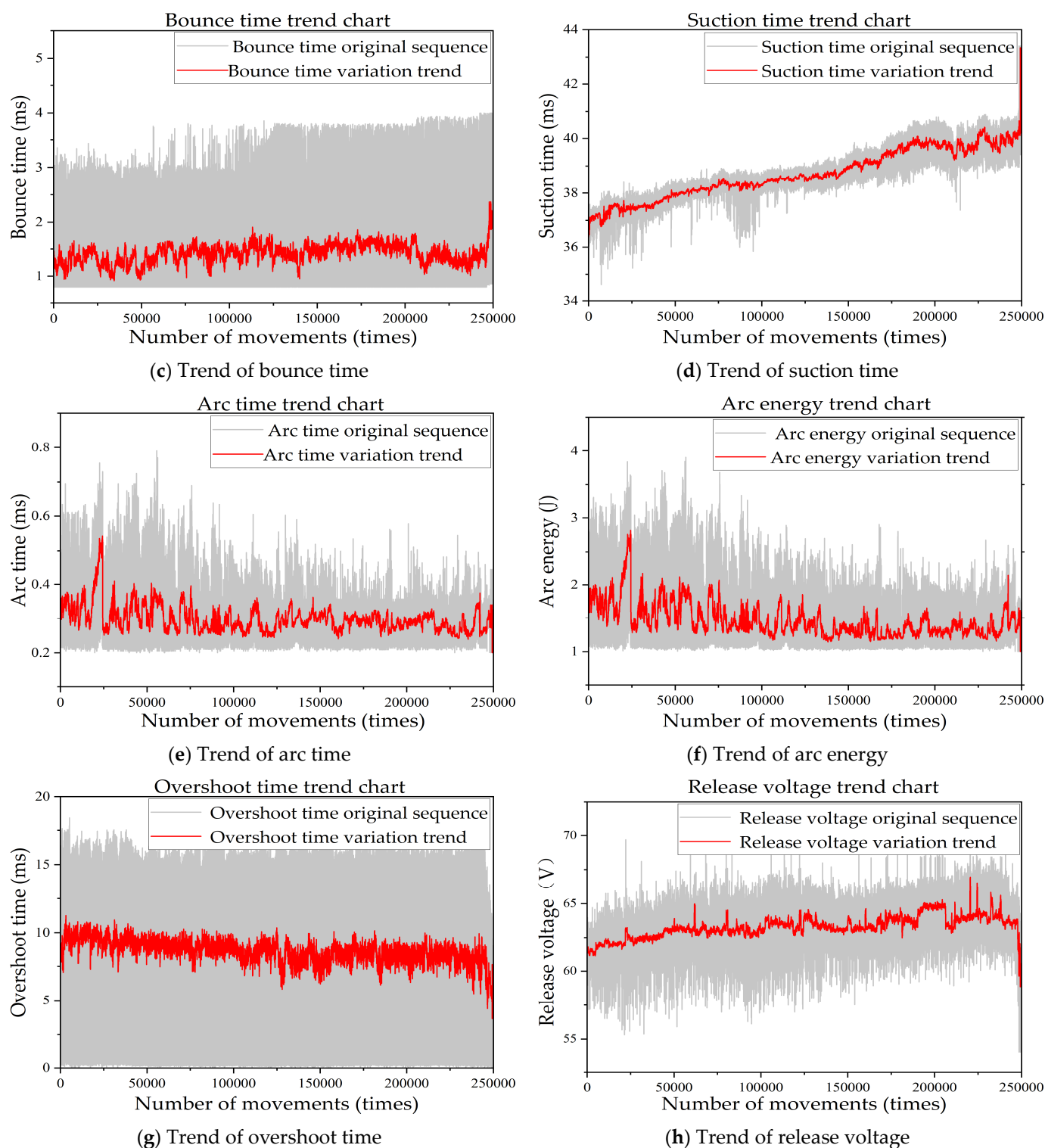


Figure 7. Trend diagram of the characteristic parameters.

4. Model Construction and Example Analysis

From the railway relay's electrical life test data, eight types of feature parameters were extracted. Using multiple feature parameters for training allows for the retention of more useful information and better representation of the railway relay's degradation process. However, having too many feature parameters can lead to some redundant features, which can increase the difficulty of model training, increase training time, and may also affect the prediction accuracy. Therefore, in this study, a dual feature selection process

was carried out using Spearman's rank correlation coefficient analysis and random forest importance analysis to eliminate redundant feature parameters while retaining the most useful feature information.

By combining the advantages of CNN and BiLSTM, the CNN BiLSTM prediction model was constructed. Firstly, data preprocessing was performed, including data normalization. The inputs were divided into three parts: training set, validation set, and test set. The remaining action count of the railway relay was used as the label for the prediction model. Then, the optimal feature subset was fed into the model. Finally, the predicted data were reverse normalized to obtain the remaining action count prediction, which was compared and analyzed against the actual remaining action count. The prediction flow structure is illustrated in Figure 8.

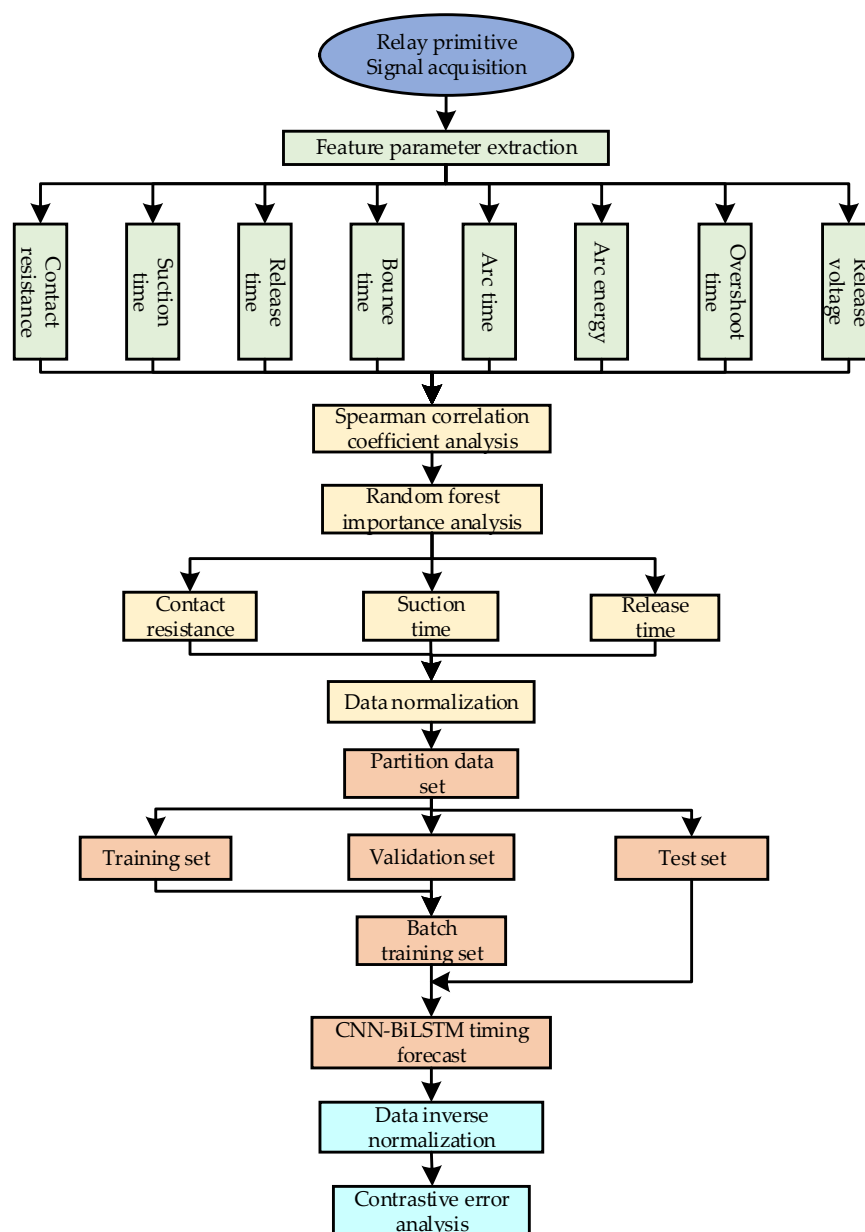


Figure 8. Structural diagram of the prediction flow.

The constructed CNN BiLSTM combined model possesses both the feature extraction capability of CNN and the stronger representation learning ability of BiLSTM. Additionally, it benefits from BiLSTM's capability to capture bidirectional time sequences, enabling better capturing of the correlation between input feature data in both directions and learning of long-term dependencies. This allows the model to more effectively predict the remaining electrical life of railway relays.

4.1. Feature Parameter Selection

The feature selection process was completed through the use of Spearman correlation coefficient analysis and random forest importance analysis. The results of the Spearman's rank correlation coefficient analysis are shown in Figure 9.

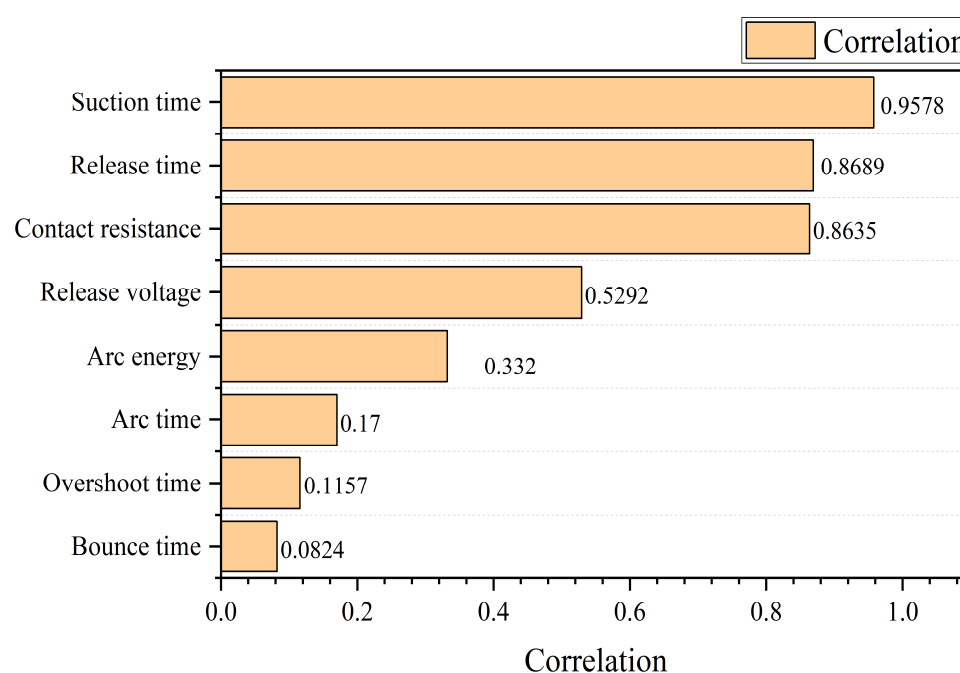


Figure 9. Spearman correlation coefficient diagram.

From the above figure, it can be observed that the Spearman's rank correlation coefficients for bounce time, arcing time, and overshoot time are all less than 0.2, indicating no significant correlation. Therefore, these features are removed from further consideration. Subsequently, the remaining feature parameters underwent random forest importance analysis, and the results are shown in Figure 10.

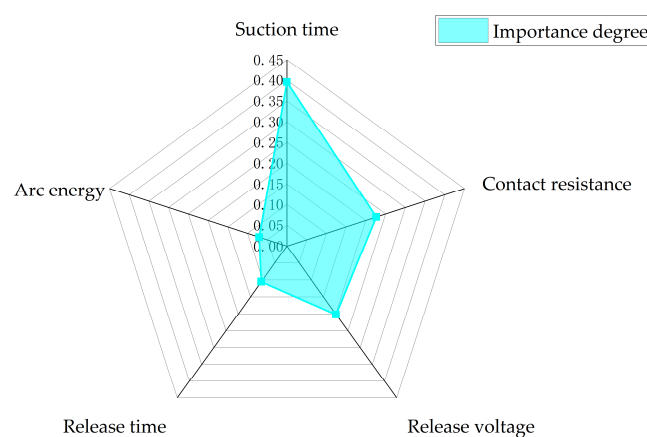


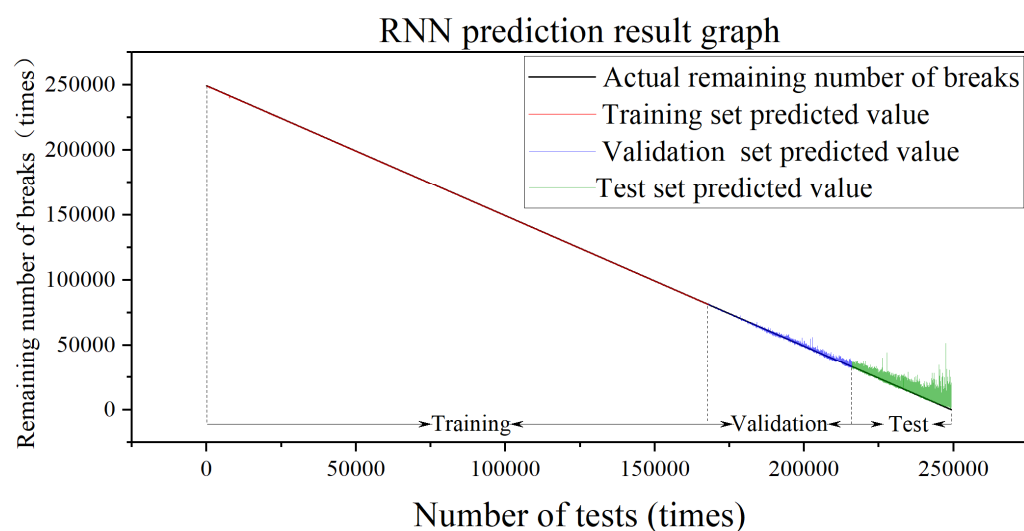
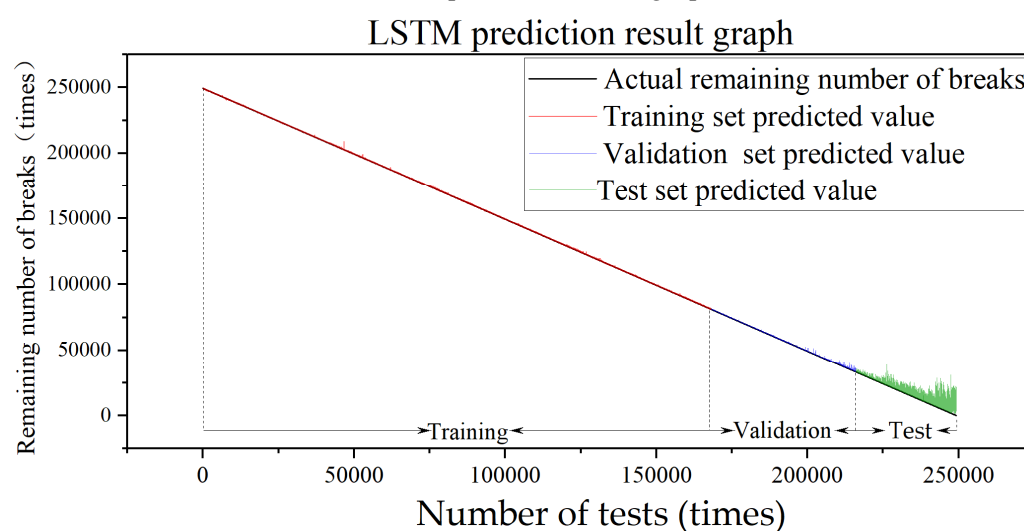
Figure 10. Random forest importance analysis chart.

Based on the results of the random forest importance analysis, it can be observed that the sum of the importance scores for the release voltage, contact resistance, and suction time (three feature parameters) is 0.824, which exceeds 0.8. This indicates that these features can effectively represent the feature information and reflect the degradation process of the relay. Therefore, they are considered as the optimal feature subset for completing the remaining life prediction process.

4.2. Comparative Analysis of the Prediction Results

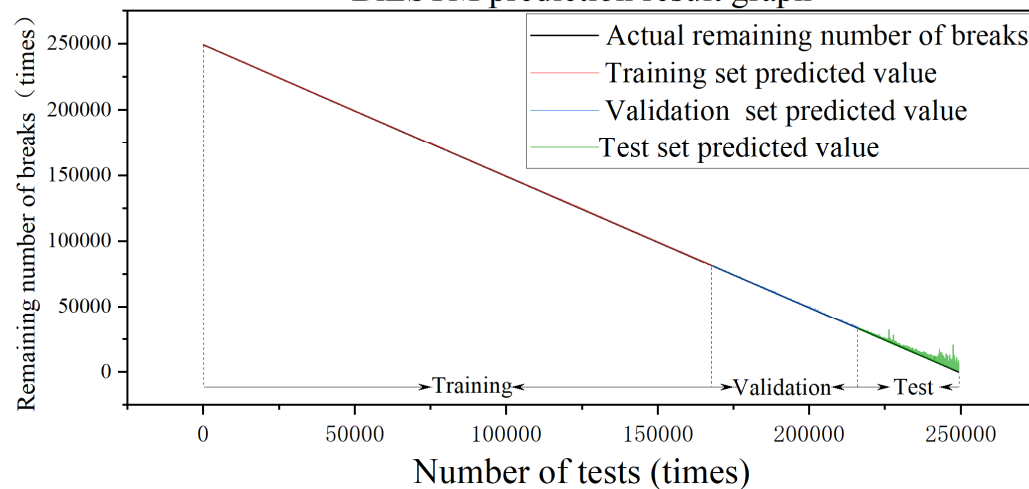
In this study, four prediction models were constructed: RNN, LSTM, BiLSTM, and CNN BiLSTM. The models were evaluated using both the original feature parameters and the feature parameter subsets obtained through dual feature selection. The prediction process was applied to estimate the remaining electrical life of railway relays. Calculate The corresponding performance evaluation indicators are calculated based on the prediction results for model prediction comparison in order to verify that the CNN BiLSTM model has more advantages compared to other models.

The comparison of prediction results before and after feature selection is shown in Figures 11 and 12, respectively. The specific evaluation metrics are presented in Tables 4 and 5.

**(a)** RNN prediction result graph

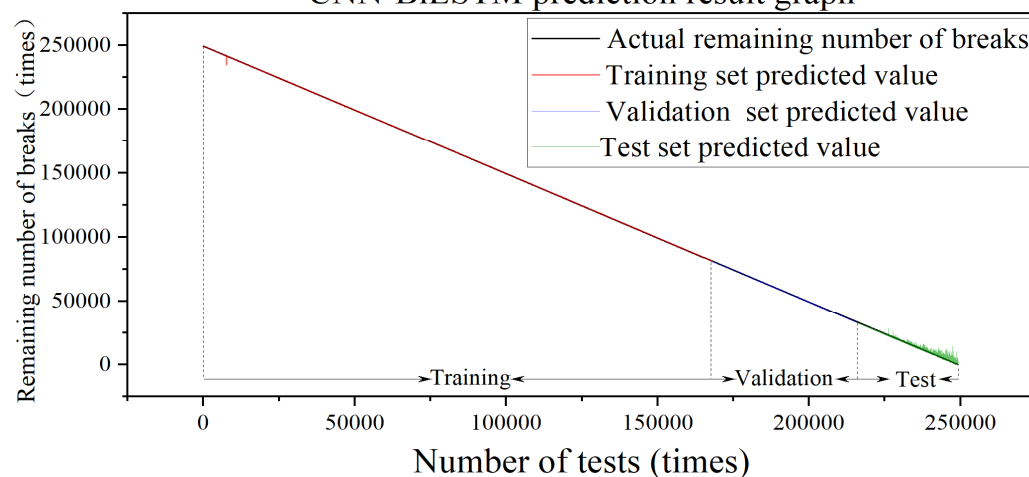
(b) LSTM prediction result graph

BiLSTM prediction result graph



(c) BiLSTM prediction result graph

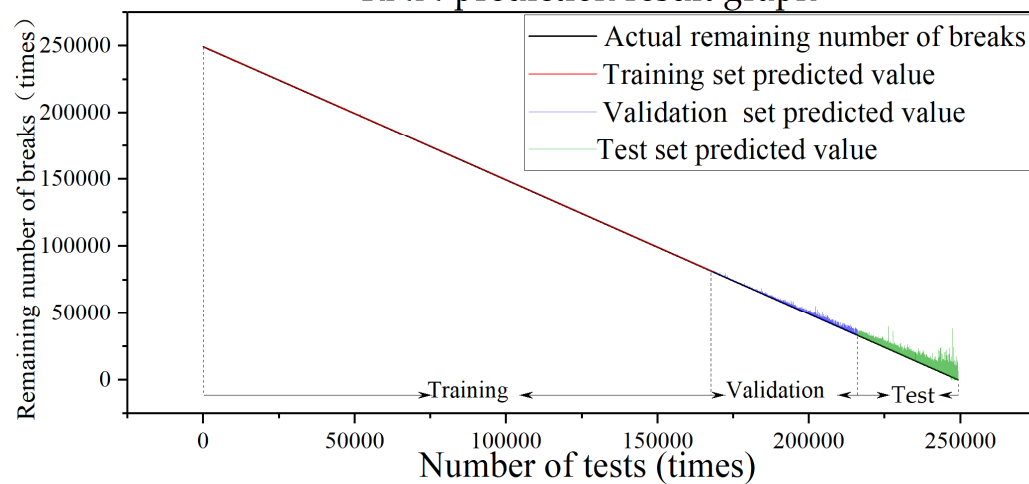
CNN-BiLSTM prediction result graph



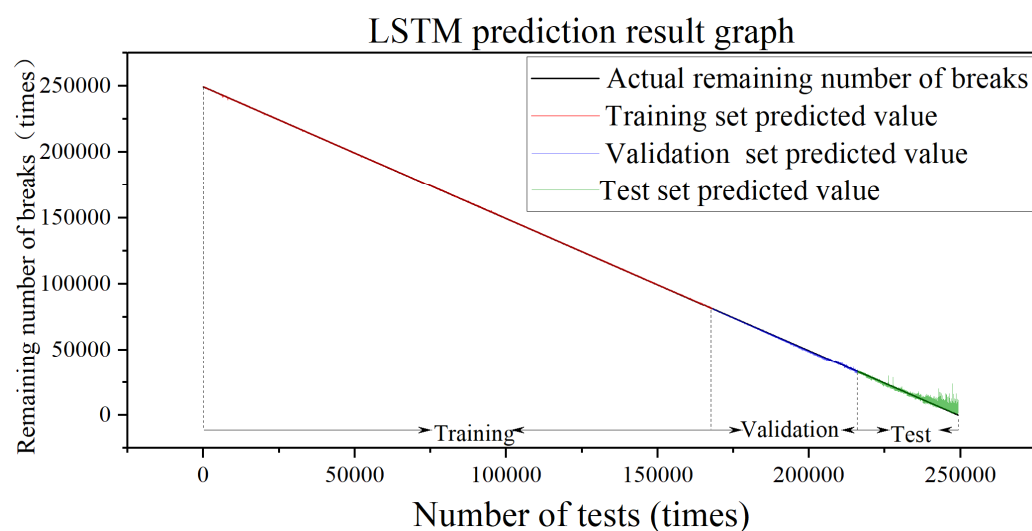
(d) CNN BiLSTM prediction result graph

Figure 11. Comparison of prediction results before feature selection.

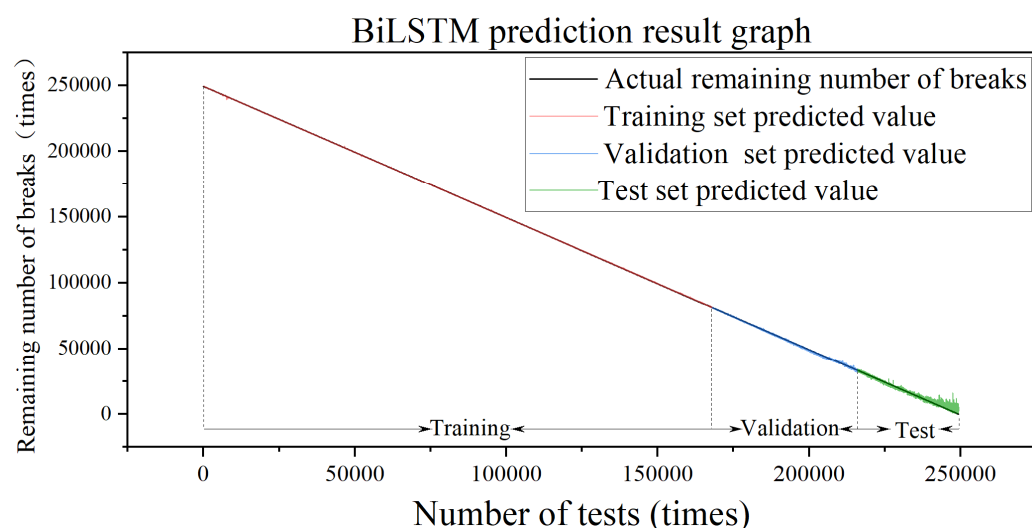
RNN prediction result graph



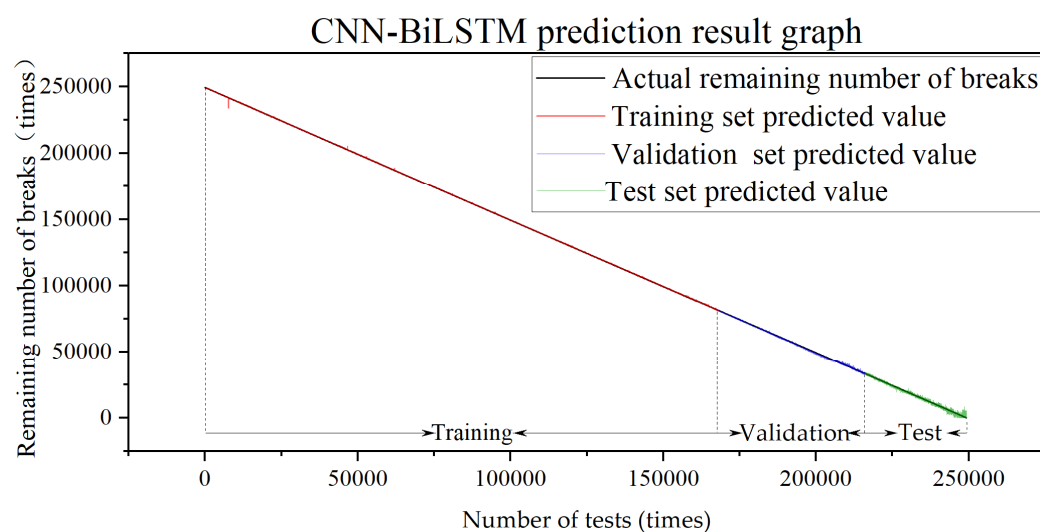
(a) RNN prediction result graph



(b) LSTM prediction result graph



(c) BiLSTM prediction result graph



(d) CNN BiLSTM prediction result graph

Figure 12. Comparison of prediction results after feature selection.

Table 4. Evaluation index without feature selection.

Evaluation Index	Featureless Selection			
Prediction model	RNN	LSTM	BiLSTM	CNN BiLSTM
<i>RMSE</i>	20,966.12	20,296.63	19,874.25	18,831.17
<i>MAE</i>	2698.24	1795.82	1045.73	711.71
R^2	0.833	0.887	0.932	0.991
Maximum error rate	0.199	0.118	0.075	0.051
Effective accuracy	80.1%	88.2%	92.5%	94.9%
Calculation time	225.7s	212.8.8s	240.5s	154.6s

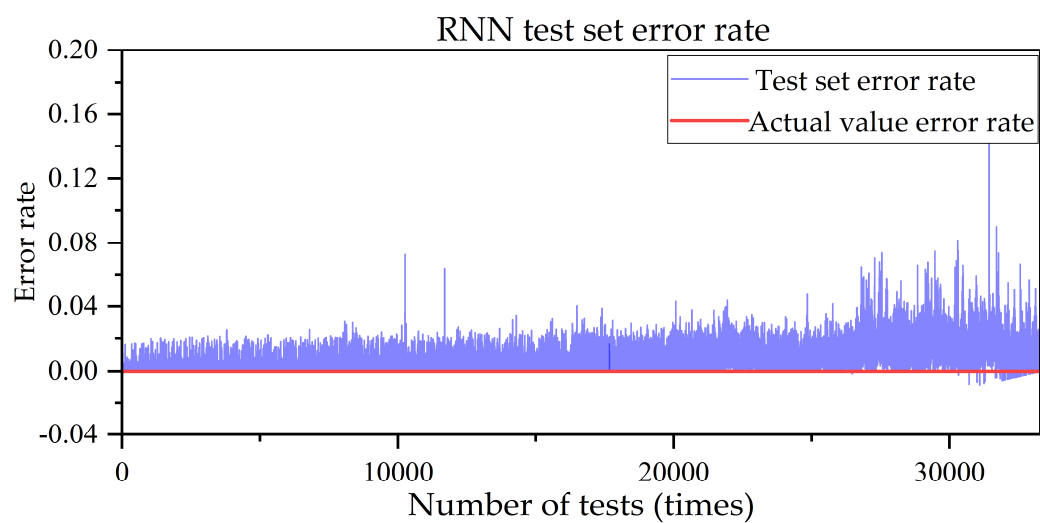
Table 5. Evaluation index with feature selection.

Evaluation Index	Feature Selection			
Prediction model	RNN	LSTM	BiLSTM	CNN BiLSTM
<i>RMSE</i>	20,395.95	19,248.89	19,068.53	18,807.37
<i>MAE</i>	1820.79	1129.14	817.34	657.28
R^2	0.92	0.963	0.979	0.994
Maximum error rate	0.145	0.088	0.067	0.037
Effective accuracy	85.5%	91.2%	93.3%	96.3%
Calculation time	72.4s	62.7s	70.6s	42.8s

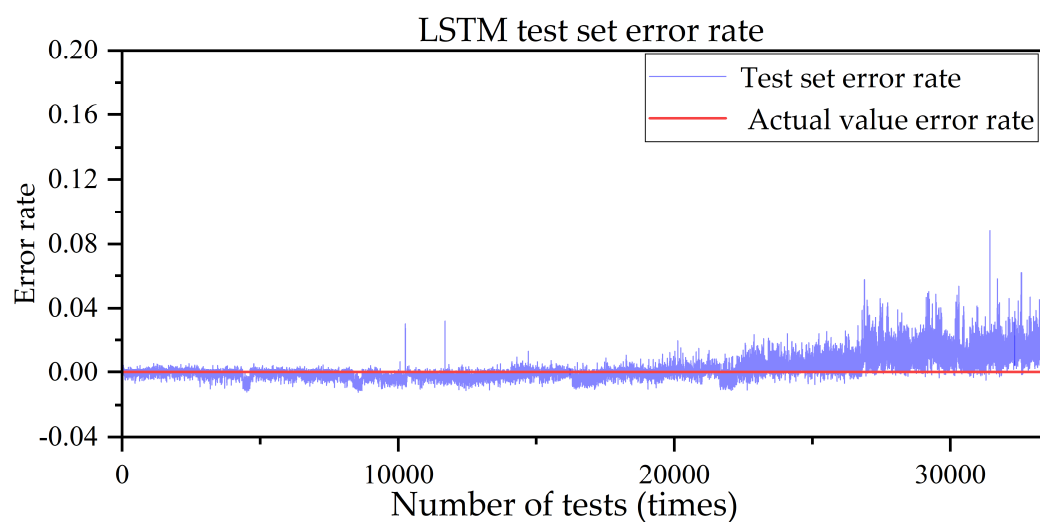
Based on the comparison of Figures 11 and 12 and the data from Tables 4 and 5, the following observations can be made:

- (1) The prediction results of the models without feature parameter selection will have some large error fluctuations, affecting the model prediction accuracy as well as stability. After the feature parameter selection, the *RMSE*, *MAE*, and maximum error rate of the four prediction models are reduced to different degrees and the effective accuracy and R^2 are higher, and the effective accuracy of the four models after feature parameter selection is improved by 5.4%, 3%, 0.8%, and 1.4%, which directly demonstrates the significance of feature parameter selection.
- (2) Regardless of whether the feature parameter selection process is performed or not, compared with the other three prediction models, the *RMSE*, *MAE*, and maximum error rate of the CNN BiLSTM model are significantly lower and the effective accuracy and R^2 are higher, which also directly demonstrates that the CNN BiLSTM model has a higher accuracy and predictive stability compared with the RNN model, LSTM model, and BiLSTM model. The overall advantage of the prediction model is more obvious. Moreover, the effective accuracy of the CNN BiLSTM model after the selection of feature parameters reaches 96.3%, which is 10.8%, 5.1%, and 3% higher than the other three prediction models, indicating that the CNN BiLSTM prediction model has more obvious advantages in the prediction of the residual electrical life of railway relays.

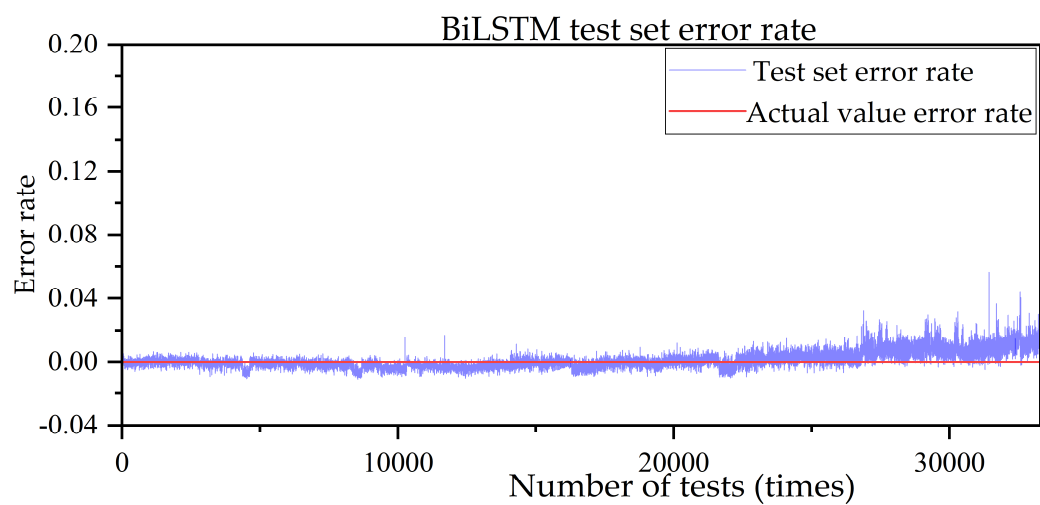
The actual error rate directly reflects the quality of the prediction results. By calculating the actual error rate using the prediction results from the test set, further comparisons among different models can be made. The prediction error rates of the four models are shown in Figure 13.



(a) RNN prediction error rate



(b) LSTM prediction error rate



(c) BiLSTM prediction error rate

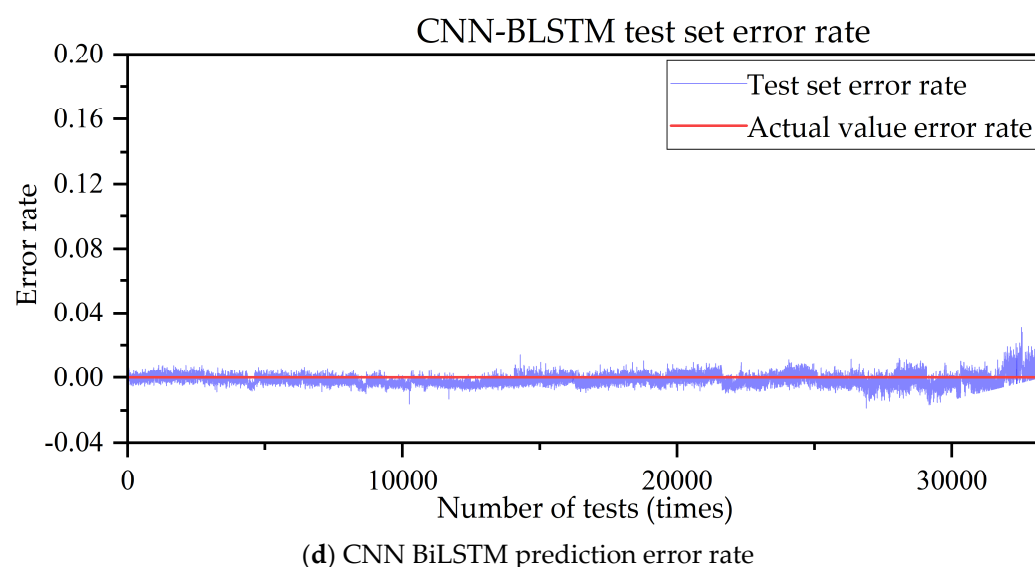


Figure 13. Comparison of diagrams of model errors.

From the comparison of the prediction error rate among the four models, it is evident that the CNN BiLSTM model has a significantly lower prediction error rate compared to the RNN, LSTM, and BiLSTM models. Moreover, the CNN BiLSTM model exhibits overall lower fluctuation in the prediction error rate, indicating better stability.

Considering the prediction results, performance evaluation metrics, and the comparison of prediction error rates on the test set, it is clear that the CNN BiLSTM model outperforms the others, with higher effective accuracy and better stability. Therefore, the CNN BiLSTM model is more suitable for railway relay remaining electrical life prediction.

5. Conclusions

In this study, we proposed a railway relay remaining electrical life prediction method based on CNN BiLSTM, utilizing the extracted feature parameters from the railway relay electrical life test to model the degradation process. The following conclusions were drawn through error analysis and model comparisons:

- (1) The process of Spearman correlation coefficient analysis and random forest importance selection directly influences the accuracy and stability of railway relay remaining electrical life prediction. Removing redundant features to obtain the optimal feature subset is an indispensable part of railway relay electrical life prediction.
- (2) By introducing deep learning prediction methods into the relay electrical life prediction process, we constructed the CNN BiLSTM combined model. The comparison of different models' time series prediction results showed that the CNN BiLSTM prediction model achieved an effective accuracy of 96.3%, which was 10.8% higher than the RNN model, 5.1% higher than the LSTM model, and 3% higher than the BiLSTM model. Its stability and prediction accuracy were higher, making the CNN BiLSTM model more suitable for railway relay remaining electrical life prediction, and it demonstrated better performance in practical applications.
- (3) This method provides a new approach for railway relay life prediction, which is beneficial for relay maintenance and can improve economic efficiency, reduce failure rates, and enhance safety and stability, thereby ensuring the safe and stable operation of the entire train.
- (4) This study completed the test process based on the relay life test regulations. However, there are still slight differences between the test process and the actual working conditions of the relay. For example, the presence of vibration, humidity, and high temperature conditions in actual working conditions can affect the service life of the

relay. Therefore, in the next step, we will conduct railway relay life prediction tests in a real working environment to further verify the feasibility of this method.

Author Contributions: S.L. and Y.L. conceived and completed the writing of the manuscript. S.G. and C.X. finished debugging the program and running the results. J.L. and Y.C. performed experimental analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Major Science and Technology Project of Liaoning Province (2020 JH 1/10100012); Project of Liaoning Education Department (LJGD2020001); and Shenyang Young and Middle-aged Science and Technology Innovation Talents Program (RC210354).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available because the supporting project has a confidentiality agreement.

Acknowledgments: Thank you to the editor and reviewers for their valuable comments and suggestions, which helped us to improve the manuscript.

Conflicts of Interest: The authors declare no competing financial interest.

References

- Li, W.H.; Ma, S.N.; Shen, P.G.; Wang, N. Research on life prediction of railroad relays under vibration conditions. *J. Electr. Eng.* **2017**, *12*, 9–10. <https://doi.org/10.11985/2017.07.002>.
- Li, W.H.; Yin, S.Z.; Zhao, Z.Y.; Xia, Y.H.; Pan, R.Z. Research on the fusion method of railroad relay electromechanical parameters based on adaptive variable weights. *J. Railw. Sci. Eng.* **2022**, *19*, 249–251. <https://doi.org/10.19713/j.cnki.43-1423/u.T20210063>.
- Meng, J.B. Analysis of common failure causes of contactors in railroad coaches. *Railw. Veh.* **2022**, *60*, 124–126. <https://doi.org/10.3969/j.issn.1002-7602.2022.04.027>.
- Gong, X.; Tong, T.; Wang, Y.H. Failure analysis and optimized maintenance strategy discussion of B series European standard relay on board of subway signal system. *Electromechanical Inf.* **2023**, *10*, 72–75. <https://doi.org/10.19514/j.cnki.cn32-1628/tm.2023.10.020>.
- Guan, X.; Lu, Z.G.; Yue, B.Q.; Zhang, W.J. Study on life prediction of electromagnetic relays based on weight optimization combination model. *Electr. Appl. Energy Effic. Manag. Technol.* **2022**, *6*, 45–50. <https://doi.org/10.16628/j.cnki.2095-8188.2022.06.009>.
- Li, L.L.; Zhang, S.N.; Li, Z.G.; He, P.J. Relay life prediction method based on rough set theory and life initial state information. *J. Electrotechnol.* **2016**, *31*, 46–53. <https://doi.org/10.19595/j.cnki.1000-6753.tces.2016.18.006>.
- Li, Z.G.; Liu, B.Y.; Li, L.L.; Sun, D.W. Relay life prediction based on wavelet packet transform and RBF neural network. *J. Electrotechnol.* **2015**, *30*, 233–240. <https://doi.org/10.19595/j.cnki.1000-6753.tces.2015.14.032>.
- Li, H.; Sun, D.W.; He, P.J.; Li, L.L. Relay life prediction method based on overrange time regression model. *J. Electrotechnol.* **2013**, *28*, 414–423. <https://doi.org/10.19595/j.cnki.1000-6753.tces.2013.s2.077>.
- Qiao, Q.Y.; Wang, Z.B.; Liu, B.X.; Chen, K.N. Storage life prediction of electromagnetic relay based on WOA-SVM model. *Electr. Appl. Energy Effic. Manag. Technol.* **2023**, *2*, 1–10. <https://doi.org/10.16628/j.cnki.2095-8188.2023.02.001>.
- Zhang, T.W.; Guang, L.; Liu, D.; Shi, J.T.; Yang, Y.Z. Analysis of common failure causes and failure mechanism of electromagnetic relays. *Intern. Combust. Engines Accessories* **2023**, *13*, 68–70. <https://doi.org/10.19475/j.cnki.issn1674-957x.2023.13.014>.
- Wang, X.; Zhao, X.H.; Yao, W.; Sun, H.D. Railroad relay life test system based on PIC microcontroller. *Railw. Commun. Signal* **2023**, *59*, 86–90. <https://doi.org/10.13879/j.issn.1000-7458.2023-02.22322>.
- Lan, C. Research on life prediction method of automobile electromagnetic relay based on double convolutional neural network. *Northeast For. Univ.* **2021**, *8*, 1–44. <https://doi.org/10.27009/d.cnki.gdblu.2021.000404>.
- Li, J.X.; Liu, Y.; Pei, C.X.; Li, N. Research on relay life assessment technology for CRH380B series rolling stock. *China Test* **2020**, *46*, 82–89+118. <https://doi.org/10.11857/j.issn.1674-5124.2019090030>.
- Li, D.H.; Zhou, X.; Wang, A.; Wang, R.; Zhai, G.F. Batch Vacuum AC Contactor Mechanical Life Prediction Considering Manufacturing Parameters. *Trans. China Electrotech. Soc.* **2023**, *38*, 1982–1990. <https://doi.org/10.19595/j.cnki.1000-6753.tces.211894>.
- Li, Z.Z. Research on life prediction method of electromagnetic relay based on improved DBN and Softmax regression. *Northeast For. Univ.* **2020**, *2*, 1–47. <https://doi.org/10.27009/d.cnki.gdblu.2020.000697>.
- Li, Y.B. Research on life prediction model of balanced force electromagnetic relay based on failure mechanism. *Harbin Inst. Technol.* **2022**, *6*, 1–54. <https://doi.org/10.27061/d.cnki.ghgdu.2022.004332>.
- Yazdanxin, P.; Sharifan, S. A multiscale ensemble of LSTM/GAN deep learning architecture for multistepahead cloud workload prediction. *J. Supercomput.* **2021**, *77*, 11052–11082. <https://doi.org/10.1007/s11227-021-03723-6>.
- Zhou, Y.Q.; Kumar, A.; Gandhi, C.P.; Vashishtha, G.; Tang, H.; Kundu, P.; Singh, M.; Xiang, J. Discrete entropybased health indicator and LSTM for the forecasting of bearing health. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 120–137. <https://doi.org/10.1007/s40430-023-04042-y>.

19. Zhao, W.; Yuan, X.L.; Zhou, Y.X.; Li, P.Q.; Lei, E.T. A capacity allocation method for a gradient battery energy storage system considering the economic optimization over the operating life. *Power Syst. Prot. Control* **2021**, *49*, 16–24. <https://doi.org/10.19783/j.cnki.pspc.201233>.
20. Li, J.J.; Chen, J.Y.; Huang, K.Y. Multi-class power theft detection based on CNN-LSTM hybrid model. *J. Electr. Power* **2022**, *37*, 367–373. <https://doi.org/10.19781/j.issn.1673-9140.2023.01.026>.
21. Li, Z.; Ye, L.; Dai, B.H.; Yu, Y.J.; Luo, Y.D.; Song, X.R. Ultra-short-term wind power prediction method based on IDSCNN-AM-LSTM combined neural network. *High Volt. Technol.* **2022**, *48*, 2117–2127. <https://doi.org/10.13336/j.1003-6520.hve.20210557>.
22. Liu, S.X.; Gao, S.Z.; Liu, Y.; Li, J.; Cao, Y.D. Residual life prediction of AC contactors based on LSTM. *High Volt. Technol.* **2022**, *48*, 3210–3220. <https://doi.org/10.13336/j.1003-6520.hve.20220082>.
23. Ru, H.Z.; Ding, Y.F.; Wang, P.Y. Ultrashort-term wind power prediction based on information fusion and stacking model. *Renew. Energy* **2022**, *40*, 344–349. <https://doi.org/10.13941/j.cnki.21-1469/tk.2022.03.006>.
24. Yang, J.; Liu, Y.; Huang, K.P.; Luo, Y.Z.; Niu, S.Q.; Wang, W.; Huang, J.F.; Zhang, L.; Li, P.; Li, H.W. A method for estimating the available generation power of wind farms considering generation conditions and station losses. *Power Gener. Technol.* **2023**, *44*, 235–243. <https://doi.org/10.12096/j.2096-4528.pgt.21135>.
25. Zhou, J.L.; Zhou, Z.; Cui, L. Classification of loudspeaker anomalous sound by applying variational modal decomposition and random forest feature selection algorithm. *Vib. Shock* **2022**, *41*, 277–283. <https://doi.org/10.13465/j.cnki.Jvs.2022.20.034>.
26. Xue, Y.; Wang, L.; Wang, S.; Zhang, Y.F.; Zhang, N. An ultra-short-term wind power prediction model combining CNN and GRU networks. *Renew. Energy* **2019**, *37*, 456–462. <https://doi.org/10.13941/j.cnki.21-1469/tk.2019.03.023>.
27. Pereira, J.; Sarayva, F. Convolutional neural network applied to detect electricity theft: A comparative study on unbalanced data handling techniques. *Int. J. Electr. Power Energy Syst.* **2021**, *131*, 107085. <https://doi.org/10.1016/j.ijepes.2021.107085>.
28. Zhu, L.J.; Xun, Z.H.; Wang, Y.X.; Cui, Q.; Chen, W.Y.; Lou, J.C. Short-term power load forecasting based on CNN-BiLSTM. *Grid Technol.* **2021**, *45*, 4532–4539. <https://doi.org/10.13335/j.1000-3673.pst.2021.0470>.
29. Ma, X.L.; Tao, Z.M.; Wang, Y.H.; Yu, H.Y.; Wang, Y.P. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp. Res. Part C Emerg. Technol.* **2015**, *54*, 187–197. <https://doi.org/10.1016/j.trc.2015.03.014>.
30. Shi, J.R.; Zhao, D.M.; Wang, L.H.; Jiang, T.X. Short-term wind power prediction based on RR-VMD-LSTM. *Power Syst. Prot. Control* **2021**, *49*, 63–70. <https://doi.org/10.19783/j.cnki.pspc.210123>.
31. Yang, L.; Wu, H.B.; Ding, M.; Bi, R. Short-term load forecasting for Bi-LSTM networks with feature selection in new energy grids. *Power Syst. Autom.* **2021**, *45*, 166–173. <https://doi.org/10.7500/AEPS20200202002>.
32. Ren, J.J.; Wei, H.H.; Zhuo, Z.L.; Hou, T.T.; Yuan, Y.L.; Shen, J.Q.; Wang, X.M. Ultra-short-term power load forecasting based on CNN-BiLSTM-Attention. *Power Syst. Prot. Control* **2022**, *50*, 108–116. <https://doi.org/10.19783/j.cnki.pspc.211187>.
33. Yin, J.; Liu, B.; Sun, G.B.; Qian, X.W. Remaining life prediction of lithium-ion batteries based on migration learning and DAE-LSTM. *J. Electrotechnol.* **2023**, *9*, 1–14. <https://doi.org/10.19595/j.cnki.1000-6753.tces.221890>.
34. Zhou, J.Q.; Zhu, W.C. WH. Rolling Bearing Life Prediction Based on DTW Similarity and Bi-LSTM. *Mach. Tool Hydraul.* **2022**, *50*, 179–184. <https://doi.org/10.3969/j.issn.1001-3881.2022.003>.
35. GB/T 14048.2-2020; Low-Voltage Switchgear and Controlgear. State Administration for Market Supervision and Administration: Beijing, China, 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.