A novel health state prediction approach based on artificial intelligence combination strategy for compensation capacitors in track circuit

Conghui Wang[1] · Shiwu Yang[1] · Chang Liu[2]

Abstract

The health management of railway signal equipment in the high-speed railway is a key link between intelligent operation and maintenance. Accurately predicting the health state of compensation capacitors is of great significance to ensure the reliable work of track circuits. This paper proposes an improved deep neural network algorithm focusing on the problem of long-term accurate health forecasts for compensation capacitors. First, establishing a transmission state model for degradation mechanism mining, the difference function that can quantitatively evaluate features is defined by piecewise processing cab signaling receiving voltage. Introducing the degradation model, predictive driving under both model and data is implemented. Then, the convolutional neural networks and bidirectional long–short-term memory are combined and improved to construct a novel artificial intelligence combination strategy, while parameters are optimized based on the sparrow search algorithm. Finally, facing the conditional repair of compensation capacitors, we set a reasonable threshold for the occurrence of hidden dangers to complete fault warning. This novel and practical approach effectively explores the procedure of prognosis and health management, while the refined maintenance will better utilize current monitoring information, helping the intelligence and accuracy of safety control decision-making.

Keywords Track circuit · Health state prediction · Compensation capacitor · Deep neural network · Sparrow search algorithm (SSA)

1 Introduction

To ensure the safe and efficient operation of the high-speed railway, the signal system is upgrading from fault repair to the frontier field of intelligent operation and maintenance [1]. The ZPW-2000 track circuit is the ground infrastructure

---

Extended author information available on the last page of the article

equipment of the Chinese train control system (CTCS), in which the compensation capacitor is an important outdoor component for track circuit signal to transmit stably in rails. Due to their huge quantity, scattered installation, and vulnerability to environmental impact, the failure rate of compensation capacitors is remarkably higher than other equipment, bringing difficulties in field maintenance. The failure may lead to the decreased transmission quality of track circuits, resulting in false occupancy indication or temporary degradation in the occurrence of onboard cab signals even threatening both railway operation and personal safety. The research on the fault prediction of compensation capacitors is the core task to improve track circuit safety, which helps to find equipment emergencies and hidden dangers. However, the maintenance of compensation capacitors does not meet the requirement of conditional repair, and the prognosis of its health state needs to be resolved urgently.

Many scholars have researched the health management issues of compensation capacitors from the perspectives of fault diagnosis and capacitance estimation in recent years. Debiolles proposed a method of combining partial least squares regression and the neural network to achieve failure monitoring of compensation capacitors [2]. They also established a transfer confidence model [3]. Oukhellou proposed the combination of empirical modal decomposition and Hilbert transformation to achieve fault monitoring of compensation capacitors [4]. Oukhellou utilized pattern recognition and information fusion technology to propose a failure detection method for compensation capacitors based on the Dempster–Shafer belief function [5]. Xu established a mathematical model of the cab signaling-induced voltage amplitude envelopes and achieved compensation capacitor failure diagnosis [6]. Yang established a deep hybrid nuclear network combination gated recurrent unit (GRU) model identification fault location and achieved compensation capacitor fault positioning [7]. Feng constructed the valuation function to realize the online estimate of compensation capacitors [8]. Wang established a finite element model and proposed a regression calculation method for compensation capacitors [9]. However, limited by the difficulties of constructing an accurate physics-of-failure model to describe the performance degradation properly, a health prognosis focusing on the compensating capacitor has yet to be studied. On the one hand, the railway signaling system can be defined as a complex one with diversity, correlation, and integrity features, including the interaction/coupling between numerous subunits and units, causing difficulty in exploring the spatiotemporal evolution law of component states. On the other hand, in safety–critical compensation capacitor systems, there are various fault modes under high-speed railway scenarios and complex interference conditions, presenting obvious limitations in quantitatively characterizing the relationship between data monitoring and degradation levels. Noticeably, health state prediction is closely related to degradation mechanisms. Learning and acquiring failure knowledge is an effective way to deeply understand the causes and phenomena of degradation, thereby implementing proactive prevention.

The quantitative methods for degradation mechanisms are mainly divided into model and data driven, exploring the spatiotemporal evolution of component states in complex dynamic scenarios [10]. Because the model usually needs to utilize the

knowledge of time-varying loading conditions, environmental factors, internal structure, material properties, and failure mechanisms in the life cycle of compensation capacitors. The knowledge is practically difficult to describe by mathematical equations. The construction of the model may resort to multiple professional fields and complicated experiments, which are impractical and complex. Furthermore, since the signaling risk sources at the railway scenes are complex, dynamic, and diverse, historical fault or degradation data with quality and quantity standards suitable for reasonable estimation and prediction are arduous to collect under usual circumstances, which inevitably affects feature extraction. Therefore, we expect to describe the degradation mechanism of compensation capacitors through a novel method to assist in extracting degradation knowledge, further changing the current maintenance lag. Even when the forecasted data volume insufficiently covers an entire cycle, health indicators can be extracted from knowledge to predict future evolution.

In the field of fault prediction for railway equipment, there have been some methods in recent years to effectively prevent railway safety incidents or accidents caused by personnel factors, geographical environment (including geological and climatic aspects), and equipment quality. Hu proposed the method for failure prognosis of the HVAP track circuit based on gray theory and expert systems [11]. Simone used long short-term memory (LSTM) deep learning algorithms to achieve predictive maintenance of railway rolling stock equipment [12]. Kang established an online abnormal perception model based on the LSTM network to accurately forecast the automatic train protection (ATP) failure state in the high-speed railway [13]. Nevertheless, caused by equipment and system-level degradation, maintenance interventions planned before performance degradation or failure that we focus on fundamentally differ from fault occurrence probability and prediction after fault classification discussed above. Quantitative evaluation for the characteristics of compensation capacitors is complex, owing to the influence of multi-source heterogeneous factors. Compensation capacitor degradation manifests as heterogeneity, possessing diverse degradation rules located in different spatial locations. The degradation feature may be manifested by the superposition of multiple heterogeneous compensation capacitors, while the redundant feature of ballast resistance interferes with the understanding of degradation laws. With the summary of transmission states and degradation rules, the refined and condition-based maintenance by automatically learning and extracting knowledge from massive data and heterogeneous information is challenging.

The degradation characteristics of compensating capacitors exhibit nonlinearity, which makes traditional prediction methods (such as regression methods and gray models) difficult to accurately describe and predict. Neural networks (NN) have advantages in nonlinear mapping [14]. Shi proposed an improved LSTM network for predicting ionospheric parameters [15]. Liao predicts the destination of taxi passengers based on BiLSTM [16]. Rubasinghe proposed CNN and LSTM to achieve long-term load forecasting [17]. However, LSTM suffers from insufficient global information learning from historical data and neglects the correlation between before and after time. Meanwhile, faced with long-term sequences, single CNN and BiLSTM networks still present problems of feature information loss, data structure information disorders, and insufficient feature recognition.

Concerning the issue, this paper proposes a novel prediction framework based on an artificial intelligence combination strategy, which not only includes a bidirectional structure of the BiLSTM neural network to learn the forward and backward temporal relationships of sequences but also utilizes the ability of CNN [18] to extract hidden features from static information based on convolutional computation, as shown in Fig. 1. This integrated algorithm driven by the combination strategy is referred to as SLCBN in this paper. In the data acquisition aspect, the TCR induction antenna receives continuous track circuit voltage, thereby obtaining CSRV with compensation capacitor characteristics. The main contributions of proposing such a framework are as follows:

1. The SLCBN framework can describe the degradation mechanism, presenting the entire process of gradual deterioration or failure over time in the form of a failure mathematical model. Although the random dynamic effect of internal chemical reactions and external environmental changes cannot be inferred, the model can be combined with transmitted state information to provide more comprehensive knowledge.
2. Based on the spatiotemporal data characteristics of compensation capacitors, combining regularization theory and the modeling mechanism of CNN-BiLSTM [19], this paper establishes an SLCBN algorithm with L2 regularization, which prevents overfitting and improves model accuracy.
3. The framework is built to infer the trend of feature time series, in which parameters including the number of neurons, initial learning rate, and L2 regularization
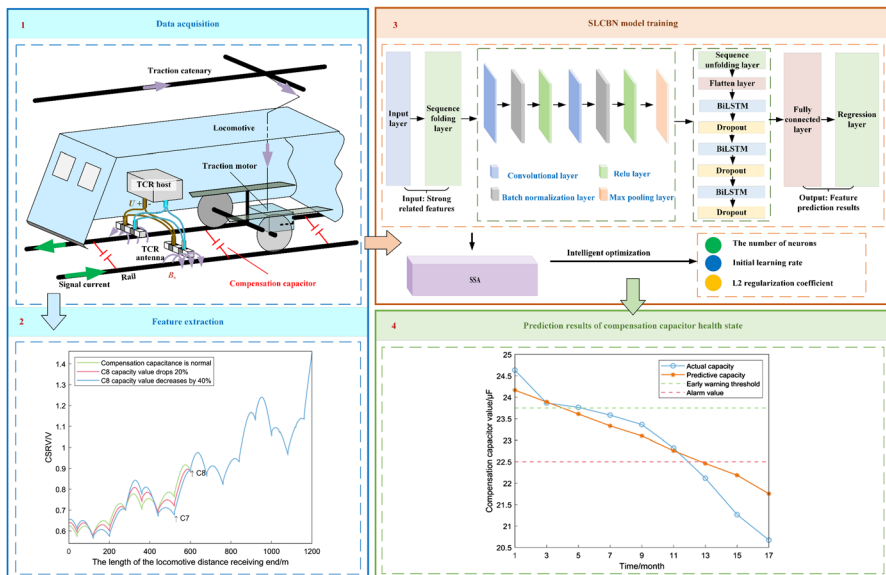
Fig. 1 A framework for predicting the health state of compensation capacitors based on a novel artificial intelligence combination strategy

coefficient are intelligently optimized by SSA. It has superior fitting, robustness, and applicability, avoiding the tedious hyperparameter tuning process.

The remainder of the paper is organized as follows: Sect. 2 constructs a degradation feature extraction strategy based on the degradation model and transmission state model of compensation capacitors. Section 3 introduces methods and processes for setting up the SLCBN model. Taking the monitoring data of China's high-speed railway field as the data source, Sect. 4 applies SLCBN to intelligently predict the health state of compensation capacitors, proving that the proposed method has significant advantages in prognosis performance by comparison and validation, and completes fault warning from the perspective of abnormal state perception by defining and calculating related parameters. Finally, Sect. 5 is the conclusion of this paper.

2 Deterioration feature extraction of compensation capacitors

Quantitative evaluation of compensation capacitor characteristics based on long-term dependence, heterogeneity, and uncertainty data is challenging. Referring to Fig. 2, compensation capacitor failure causes signal attenuation on the rail, resulting in a drop in receiving end voltage. Relying on the transmission line and four-terminal network theory, a transmission state model is constructed, attempting to reflect the health state of compensation capacitors by CSRV. Next, we integrate the compensation capacitor degradation model with state knowledge to extract the difference function as a feature, effectively grasping the degradation law of capacitors within the scope of time and space.
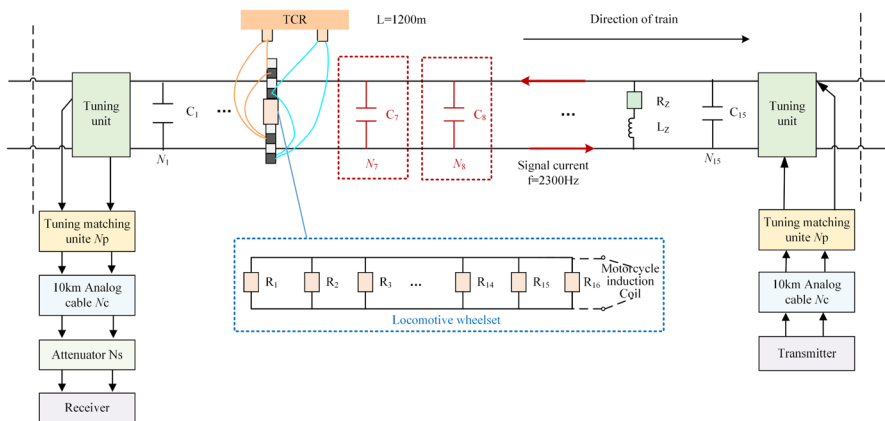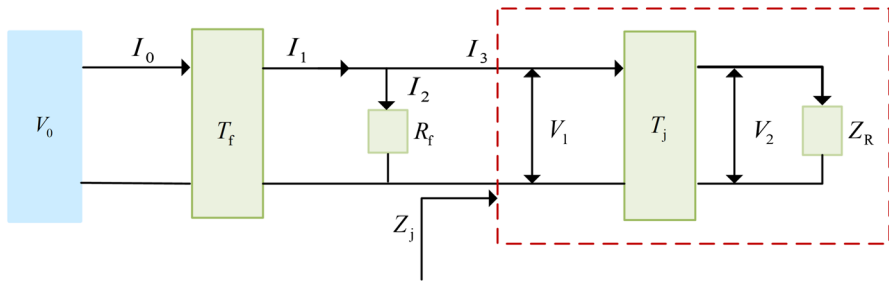

Fig. 2 The transmission state model

Table 1 The transmission matrix of each module

Category	A	B	C	D
Analog cables N_C	$\cosh(\gamma L)$	$-Z_c * \sinh(\gamma L)$	$-\sinh(\gamma L)/Z_c$	$\cosh(\gamma L)$
A matching transformer N_P(Transmitter)	n_1	$2j(\omega L/n_1 - n_1/\omega C)$	0	$1/n_1$
A matching transformer N_P(Receiver)	$1/n_2$	$2j(\omega L/n_2 - n_2/\omega C)$	0	n_2
Rails N_x	$\cosh(\gamma x)$	$-Z_c * \sinh(\gamma x)$	$-\sinh(\gamma x)$	$\cosh(\gamma x)$
Compensation capacitors $N_1 \sim N_{15}$	1	0	$j\omega C_b$	1
The shunting resistance N_{R_f}	1	$1/R_f$	0	1
The attenuator N_S	s	0	0	$1/s$

Fig. 3 Sub-equivalence four-terminal network model

2.1 Transmission state modeling

Introducing the basic structure of track circuits [20] and the transmission lines theory [21], a transmission state simulation model is established to compute the dynamic characteristics of track circuits. Further, we analyze the correlation between the dynamic degradation of compensation capacitors and CSRV, providing theoretical support for health state prognosis.

Using the four-terminal network theory to analyze track circuit transmission and distribution issues, the locomotive is usually simplified into a shunting resistance. Taking the CRH-2 electric multiple units (EMU) [22] with 16 wheelsets as an example, a simulated calculation model is given as shown in Fig. 2.

The calculation of signal transmission in track circuits can be equivalent to the cascade of four-terminal networks. Transmission matrixes are shown in Table 1.

To clarify the mapping relationship between the health state of compensation capacitors and CSRV, we establish a locomotive shunting model [23] as shown in Fig. 3.

In Fig. 3, V_0 represents the output level of the transmitter; T_f expresses a four-terminal network from the transmitter to the cab receiving end; R_f denotes the equivalent of each wheel pair; I_2 is the signal current that flows through the short-circuit wheel pair; T_j means a four-terminal network from an equivalent shunting point to the receiver terminal; Z_R represents the receiving end impedance; Z_j is the impedance from the

shunting resistance to the receiving end; and V_2 indicates residual voltage. Then, the short-circuit current of the cab signal can be calculated as:

$$I_1 = \frac{V_0}{T_{f_{11}}(x) \cdot R_f + T_{f_{12}}(x) \cdot \frac{(R_f + Z_j(x))}{Z_j(x)}} \tag{1}$$

According to railway signal maintenance rules [24], the quantitative relationship of CSRV is as follows:

$$V_1 = \frac{200}{255} I_1 \tag{2}$$

2.2 Construction of difference function based on state data

To clarify the correlation between the health state of compensation capacitors and CSRV, we introduce a transmission state model and simulate state data to analyze relevant characteristics, constructing a difference function to quantitatively evaluate the health state of capacitors. Taking the mainstream ZPW-2000A track circuit of the high-speed railway as an example, the relevant basic parameters are shown in Table 2 (as shown in Fig. 2) [25].

Since compensation capacitors and ballast resistance both significantly affect signal transmission in rail, we define a difference function to precisely estimate the health state, reducing the impact of heterogeneity factors and redundant features. CSRV is divided into several segments based on compensation capacitors, as shown in Eq. (3).

$$\begin{cases} V_{cm}^1 a \in \left[0, a_{c1}\right] \\ V_{cm}^{(i+1)} a \in \left[a_{ci}, a_{c(i+1)}\right] \end{cases} \quad i = 1, 2, \cdots, n-1 \tag{3}$$

where n is the total number of capacitors in track circuit district. The difference function of $V_{cm}^{(i)}(i = 1, 2, \cdots, n-1)$ is defined as

| Table 2 Relevant basic parameters of jointless track circuit	Relevant basic parameters	Symbol	Value
	Carrier frequency	f	2300Hz
	Total length of the track circuit	L_0	1200m
	Shunting resistance	$R_1 \sim R_{16}$	0.15Ω
	Value of compensation capacitors	C	25μF
	Output voltage level of the transmitter	V_T	170V
	Rail resistance	R_Z	2.1Ω/km
	Rail inductance	L_Z	1413μH/km

$$\begin{cases} M_{cm}^{(i)} = V_1^{(i)} - V_2^{(i)} \\ V_1^{(i)} = \begin{cases} V_{cm}^{(1)}(0) & i = 1 \\ V_{cm}^{(i)}(a_{c(i-1)}) & i \neq 1 \end{cases} \\ V_2^{(i)} = V_{cm}^{(i)}(a_{ci}) \end{cases} \quad (4)$$

where $i = 1, 2, \cdots, n$; $V_1^{(i)}$ and $V_2^{(i)}$ are the left and right boundary values, respectively.

Specifically, the main fault mode of compensation capacitors present capacity decrease [26]. We select the capacitor C8 as an example under different capacity values, i.e., 25 µF, 20 µF, and 15 µF, respectively. A simulation of three cases is shown in Fig. 4, where compensation capacitors at diverse positions exhibit various degradation characteristics, indicating significant heterogeneity. Also, a decline in the health state of C8 affects the amplitude of compensation capacitors ahead. And the most influenced point is located at C7, while CSRV located at subsequent positions of C8 almost remains unchanged. Results show that as the capacity decreases, the difference function value increases monotonously. The proposed function is adapted to heterogeneous characteristics, which can more precisely quantify the deterioration evolution.

To further verify the $M_{cm}^{(i)}$'s effectiveness of this paper, we simulate and analyze CSRV under different ballast resistances. Accordingly, the ballast resistance is set at 4 Ωkm, 12 Ωkm, and 36 Ωkm [27], respectively. From the simulation results in Fig. 4, changing ballast resistance will hardly affect difference function values. It can be seen that the proposed difference function can eliminate the impact of redundant characteristics on prediction.
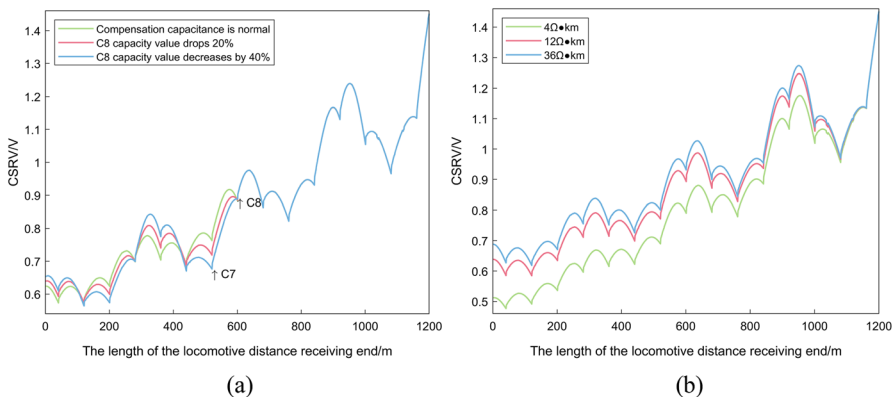
Fig. 4 The simulation results of CSRV. **a** Correlation between the health state of capacitors and CSRV; **b** correlation between ballast resistance and CSRV

2.3 Feature extraction using degradation model and data-driven approach

The data does not cover the full life cycle of compensation capacitors, which makes it difficult to provide complete support for prognosis technology. Therefore, to solve this challenging problem, we innovatively propose to combine a degradation trend model with data driven, implementing feature extraction of compensation capacitors.

China's high-speed railways adopt electrolyte capacitors [28]. Noticing that the combination of exponential and polynomial can better fit the changing trend of compensation capacitors, a dynamic degradation model is established, which is described by Eq. (5) [29]. This model is based on regression analysis of experimental data, which is sourced from NASA's Prognostics Center of Excellence (PCoE).

$$\text{Deg}(t) = \text{cap}_1 \times \exp\left(\text{cap}_2 \times t\right) + \text{cap}_3 \times t^2 - \text{cap}_4 \times t + \text{cap}_5 \tag{5}$$

where t denotes the cycle number or the time index; cap_1, cap_2, cap_3, cap_4, and cap_5 are parameters of the model, which are related to the internal impedance and aging rate of capacitors.

Based on transmission state data and a degradation trend model of compensation capacitors, we introduce the difference function to process a prediction training dataset with stubborn uncertainty factors, achieving quantitative processing for qualitative issues of invalidation mechanisms. Specifically, we employ a transmission state model to calculate CSRV, and the dynamic degradation equation of capacitors is fitted based on the standard C8 capacity of 25 μF (as shown in
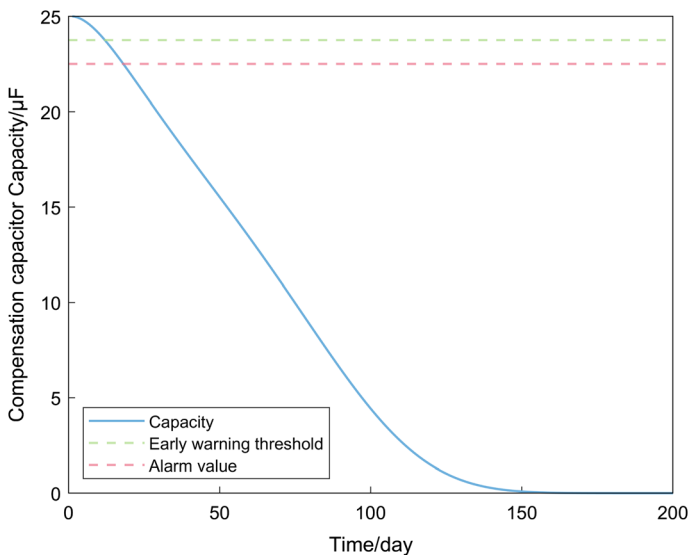

Fig. 5 Compensation capacitor capacity degradation curve

Fig. 5). The difference function of system assessment can be calculated by (4) to extract features and obtain a training set, aiming to approach changes in the health state of on-site capacitors.

3 Predictive model based on artificial intelligence combination strategy

Inspired by CNN visual space features and BiLSTM cross-sequence time information, we establish the CNN algorithmic model, select BiLSTM, and introduce SSA including the automatic parameter tuning model, completing SLCBN algorithm modeling and improvement based on artificial intelligence combination strategy.

3.1 Construction of SLCBN deep neural network

CNN is a feedforward neural network, which deeply digs out data through local connections of neurons and convolution weight sharing. It performs high-dimensional mapping, effectively reducing the number of training parameters, and improving the efficiency of feature extraction, while significantly enhancing the fitting ability of the network. The CNN model mainly consists of three layers: input layer, hidden layer, and output layer, as shown in Fig. 6.

BiLSTM can learn the normal and reverse time-sequential relationship of sequences [30]. Figure 7 reveals the structure diagram. Forgotten gate $f_2(t)$ filters and retains results of the previous memory cell. The input gate $i_2(t)$ controls the current input state. The output gate $o_2(t)$ commands the output state of the memory unit. $\tilde{c}_2(t)$ expresses the input memory cell. $c_2(t)$ denotes the output memory cell. $h_2(t)$ means the hidden state. $x(t)$ and $y(t)$ indicate the input and output at moment t, separately. σ and tan h are activation functions of sigmoid and hyperbolic tangent, respectively.

Noticeably, the CNN and BiLSTM have predominance on digging features of grid-like space and time sequence, respectively. To explore better forecast mechanisms and model training effects, we propose a novel approach toward health state prognosis
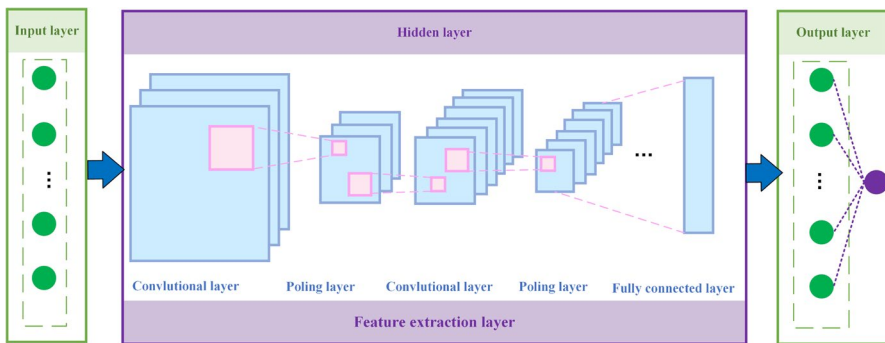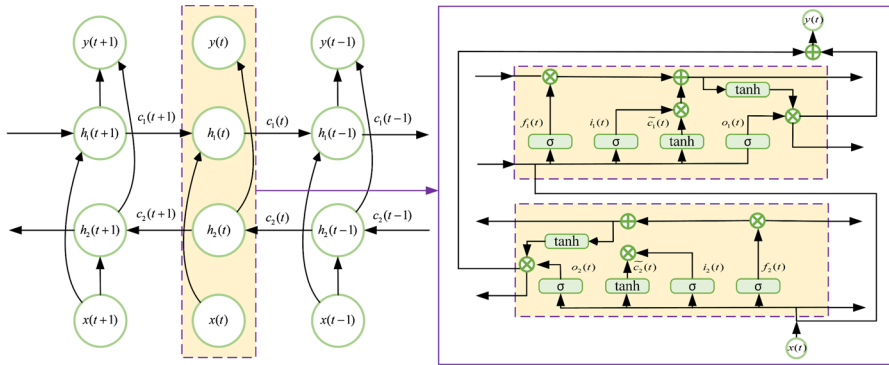

Fig. 6 Schematic diagram of CNN structure

Fig. 7 BiLSTM structure diagram

where a hybrid deep learning network consisting of the CNN, BiLSTM, and SSA, as shown in Fig. 8. Besides, different ideas have been attempted to introduce the network. Overfitting often occurs in network training, which reduces prediction accuracy. Thus, introducing L2 regularization to represent features, we normalize feature representations of natural and simulated sequences, which are generated by networks of different scales in a unified feature space. Factor matrices or variables can be regularized by adding terms to the objective function:

$$\min_X \frac{1}{2} X - T_F^2 + \lambda_X h_X(X) \tag{6}$$

in which λ_X is a hyperparameter, choosing by the user or automatically. The function $h_X(X)$ is
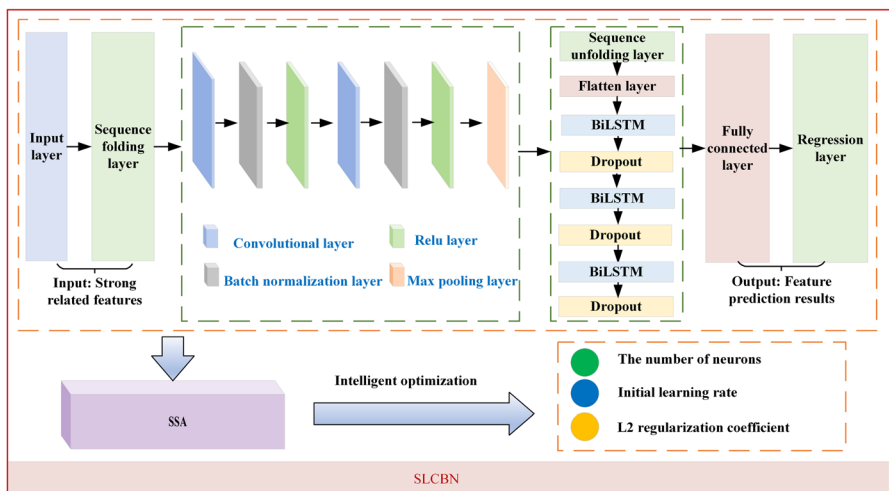

Fig. 8 SLCBN framework

$$h_X(X) = \frac{1}{2}X_2^2 \tag{7}$$

As the inability of standard CNN to process one-dimensional data, the SLCBN model constructs feature vectors for time series by converting one-dimensional raw signals into two-dimensional matrices. To reduce the impact of dimensionality between different features, we supplement the batch normalization layer on the basis of CNN while extracted features are normalized, thereby improving the accuracy of health state prediction. Furthermore, the CNN also includes two convolutional layers and a pooling layer [31], which uses the ReLU function to accelerate the convergence of the model. The SLCBN model constructs the feature representation of the time sequence through the CNN. After the feature vector is passed to the BiLSTM layer by the flatten dimensionality reduction layer, the hidden state of the degenerate feature vector will be generated, and each hidden layer is followed by a dropout layer that randomly discards some data. Finally, after three hidden layers complete the feature labeling, the output of the time series forecast tag is accomplished. For model training, an Adam optimizer with heavy decay is introduced to update the weight, and the iterative training mechanism of the model with a sliding window is used during the predicted phase, improving the accuracy of the network. The network parameter configuration is shown in Table 3.

3.2 SLCBN parameters optimization

The important parameters of the SLCBN network have a significant impact on forecast results. Traditionally, we select the number of neurons, initial learning rates, and L2 regularization coefficient according to the subject experience or grid optimization, whereas these methods may lead to overfitting or underfitting [32]. However, accurately predicting the health state of compensation capacitors is crucial for the safety of the railway system. On the one hand, since the deep integration of railway systems, misjudgment of the health state of capacitors might initiate chain reactions under other branches through knowledge and functional interactions among branches, leading to railway system equipment failures and even casualties. On the other hand, the determination of fault threshold according to the quantitative mapping of features and compensation capacitors determines its strict requirements for learning accuracy. Obviously, during algorithm designing, it is necessary

Table 3 Selection of some parameters for SLCBN network

	Layer	Setting		Layer	Setting
1	Conv1D_1	$(1\times2,10)$	6	Dropout	0.3
2	Conv1D_2	$(1\times1,10)$	7	BiLSTM	SSA
3	MaxPooling	$(1\times3,1)$	8	Dropout	0.3
4	Flatten	–	9	BiLSTM	SSA
5	BiLSTM	SSA	10	Dropout	0.3

to maximally improve the model's convergence and forecasting capability through parameter optimization [33].

As a new type of intelligent optimization algorithm, SSA is employed to optimize SLCBN network parameters, which is designed by imitating the foraging and anti-predation behavior of sparrows [34]. The optimization goal of SSA is to compensate for the forecast error, so the adaptation function is the mean squared error (MSE) of the difference function. We determine the initial conditions of SSA based on multiple tests. After debugging and performance comparison, the best parameters are obtained, and the initial conditions and results of SSA are listed in Table 4.

4 Experimental verification and application

4.1 Preprocessing of measured datasets

Evaluating the effectiveness of the proposed SLCBN framework in practical case studies of the railway field, on-site data is applied as the test set (refer to Sect. 2.3 for simulation training set acquisition) to verify the predictive performance for the algorithm proposed in this paper.

In terms of monitoring, onboard cab signals are recorded in real-time online throughout the entire process by the dynamic monitoring system (DMS) of the train control system. Consequently, CSRVs collected from the DMS are used to demonstrate the effectiveness of our proposed method. It is worth noting that voltage-related signals were collected every 2.5 m, whose typical signal characteristics are shown in Fig. 9, so the horizontal axis can directly convert into the distance between the locomotive and the track circuit receiving end. Calculated difference function values at different time points can be used as the testing set of the SLCBN model.

4.2 SLCBN model verification and comparison

Comprehensively evaluating the performance of the SLCBN algorithm for predicting the health of compensation capacitors, we use 800 sets of simulation data as the training set and 17 sets of on-site measured data to verify the performance of the model, according to four evaluation indicators, as shown in Table 5, then set the total iteration epochs as max_epochs = 500 and the batch size of each training batch_size = 256. The training set is used to reflect whether SLCBN performance

Table 4 Initial conditions and results of SSA

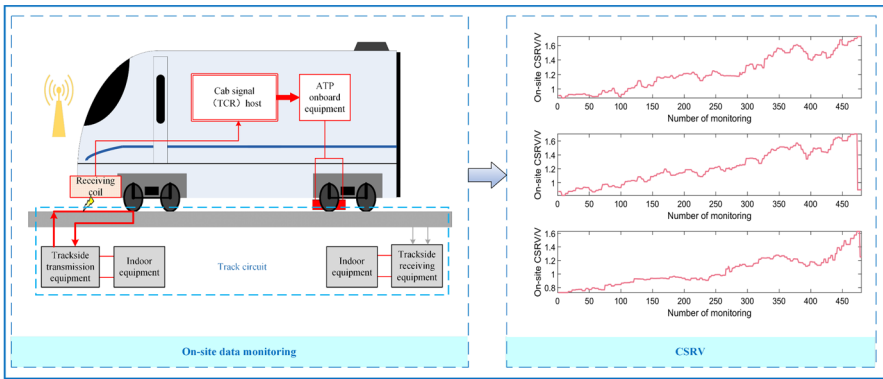| Initial conditions	Value	Optimal parameters	Optimal results
The number of sparrows S	10	The number of neurons	10
Discoverer accounts DA	20%	Initial learning rates	0.01
The maximum iteration number $iter_{max}$	5	L2 regularization coefficients	0.01
Safety threshold ST	0.8		

Fig. 9 On-site CSRV curve

exhibits positive gain during the training process, adjusting internal model hyper-parameters defined as actual iteration epochs. Specifically, by observing curves of the loss value changing with iteration, training is stopped promptly to find the best balance between fitting performance and training speed when the model tends to converge. Figure 10 evaluates the training and testing results. It can be seen from the figure that as the number of epochs increases, the fitting level of the model is remarkably improved and the neural network loss value decreases, ultimately tending to stabilize without significant overfitting or underfitting. The RMSE for training and testing is 0.018 and 0.007, respectively, while the loss values of the two (i.e., the cross-entropy loss function: $L(m, n) = -\sum m/lnq$, where n is the forecasted value and m is the true value) eventually converge nearly zero.

Comparing the fitting level of multiple prediction algorithms, i.e., SLCBN, CNN-BiLSTM, BiLSTM, LSTM, GRU [35], RNN, and CNN, training results indicate that the MAPE of the above seven algorithms is 0.03%, 0.09%, 0.2%, 0.29%, 0.3%, 0.36%, and 0.27%, respectively. Figure 11 demonstrates the performance comparison results. As we can see from Fig. 11, the SLCBN model is basically in a stable state, with four columns (i.e., red column: RMSE, blue column: MSE, green column: MAE, and purple column: MAPE) all relatively low in height, indicating that different evaluation indicators have little impact on the model and are basically in a stable state. The histograms of other models are much higher than those of SLCBN, implying that the direct modeling of capacitor degradation data affects the effectiveness of the algorithm, and this further indicates that SLCBN using regularized can mine the failure mechanism and internal knowledge of compensation capacitors, improving the stability and accuracy of our model. Noticeably, SLCBN exhibits high-precision prognosis performance with the best fitting performance, specifically RMSE = 0.00705, MSE = 0.00005, MAE = 0.00504, and MSE = 0.02608. Evidently, even in the case of a small test set, the algorithm in this paper can also forecast the difference function value with high accuracy, which has a significant advantage.

Regarding the obtained SLCBN model, we employ the 3D line chart to conveniently visualize the comparison results of each evaluation indicator under seven algorithms describing the optimal prognosis error, as shown in Fig. 12. Apparently, each

Table 5 SLCBN model evaluation indicators

Evaluating indicator	Computing formula	Interpretation
Root mean squared error (RMSE)	$\text{RMSE} = \sqrt{\dfrac{1}{n}\sum\limits_{m=1}^{n}\left(y_{pm}-y_{tm}\right)^2}$	y_{pm} – the predictive value of the m
		y_{tm} – the true value of the m
		n – the number of the test set
		RMSE can reflect the global stability of the model error
MSE	$\text{MSE} = \dfrac{1}{n}\sum\limits_{m=1}^{n}\left(y_{pm}-y_{tm}\right)^2$	MSE is used to evaluate the deviation between the measurement value and the prognosis value
Mean absolute error (MAE)	$\text{MAE} = \dfrac{1}{n}\sum\limits_{m=1}^{n}\left|y_{pm}-y_{tm}\right|$	MAE indicates the average absolute deviation of two values, which can avoid the problem of offset errors
Mean absolute percentage error (MAPE)	$\text{MAPE} = \dfrac{1}{n}\sum\limits_{m=1}^{n}\left|\dfrac{y_{pm}-y_{tm}}{y_{pm}}\right| \times 100\%$	MAPE describes a statistical indicator commonly used to evaluate the accuracy of predictions

(a)

(b)
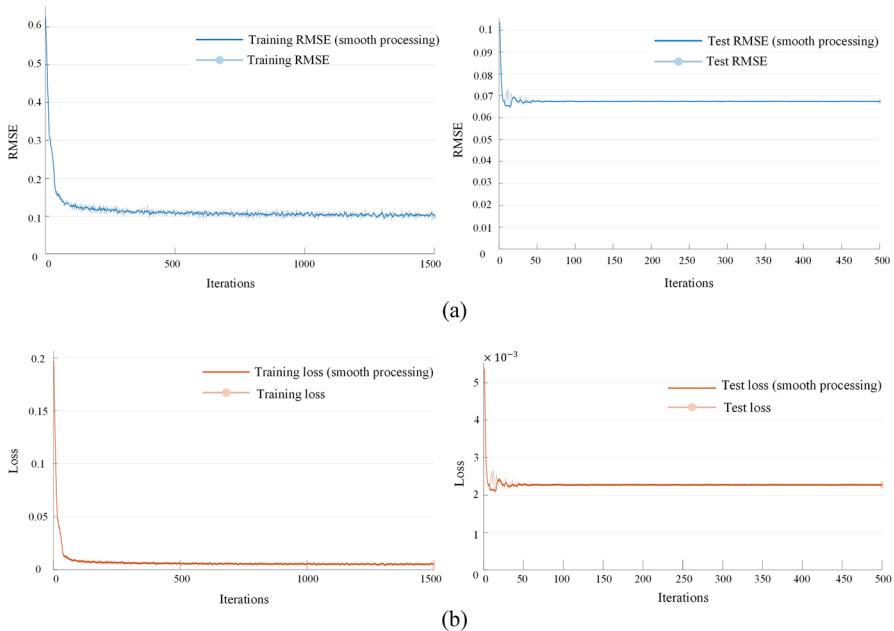
Fig. 10 Evaluation of SLCBN model results (left: training results and right: validation results). Training and validation results of **a** RMSE and **b** Loss

evaluation indicator of SLCBN is the smallest among seven algorithms, and it can be derived that the predictive algorithm 1 (i.e., dark blue curve: SLCBN) satisfies the optimal fitting ability under the evaluation by four indicators.

4.3 Application of abnormal perception

Proactively perceiving and even forecasting the potential degradation of the health state is positive for improving the safety and reliability level of railway signaling systems. Accordingly, by setting hidden danger thresholds in capacity, we aim to achieve compensation capacitor fault warning and online anomaly perception, solving the problem of delay in fault detection, diagnosis, and response.

When the depth network model predicts a 10% decrease in capacity, the alarm is triggered to achieve conditional repair. Meanwhile, an early warning is triggered when capacity decreases by 5%. We set N as the forecasted total number of compensation capacitors. n_1 indicates the number of correctly activated alarms. n_2 denotes the number of correct early warnings. Then, the accuracy of alarm and early warning is expressed as $y_1 = n_1/N \times 100\%$ and $y_2 = n_2/N \times 100\%$. In the capacity prediction aspect, the SLCBN model outputs a forecast of degradation data for capacitors in the next 17 months. Figure 13 illustrates prognosis results, indicating that among the 9-month (with a time interval of two months) predicted capacitance values using the measured dataset, there is no set of error warnings
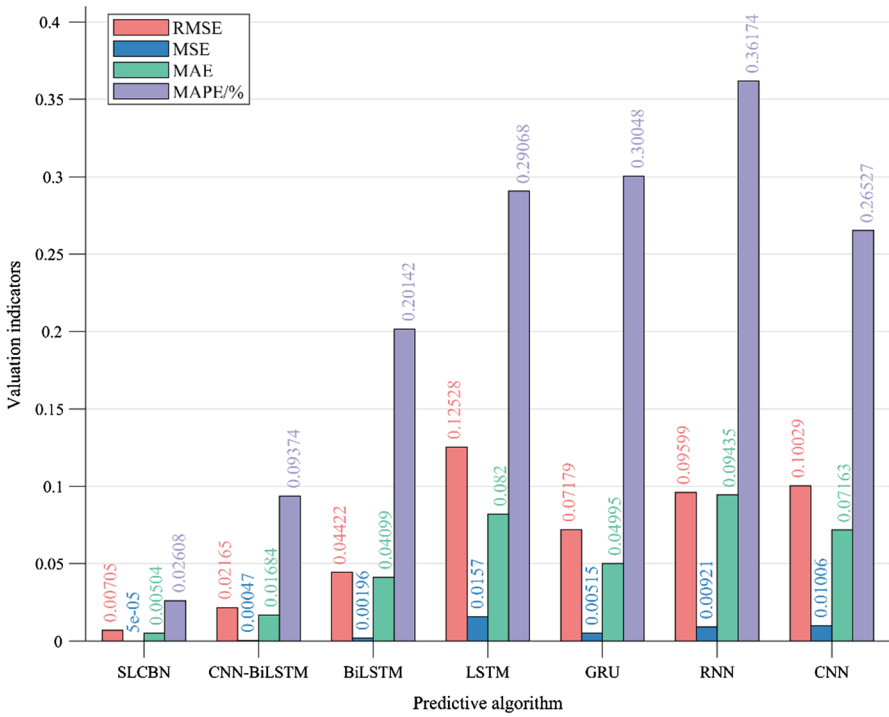
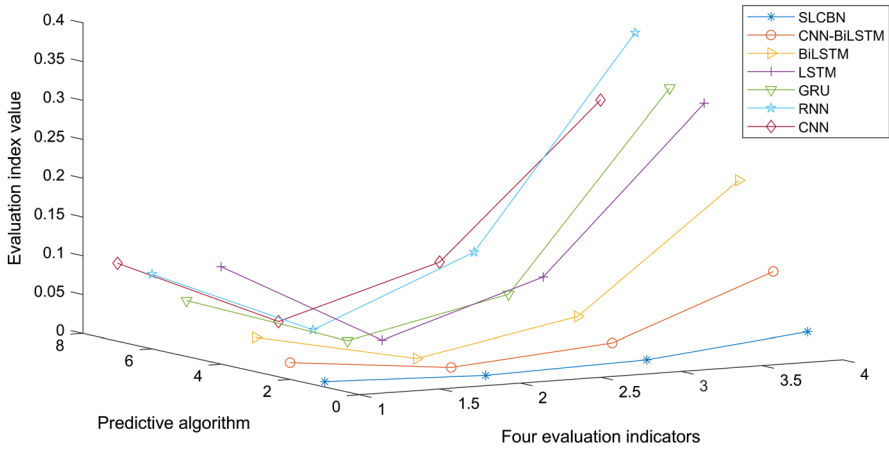**Fig. 11** Performance comparison of different prediction algorithms

Fig. 12 Algorithm performance comparison 3D line chart

indicating at the given threshold when forecasting the 5-th month warning and the 13-th month alarm, i.e., $y_1 = 100\%, y_2 = 100\%$. Although capacity prediction results have slight fluctuations, it does not affect the original intention of setting
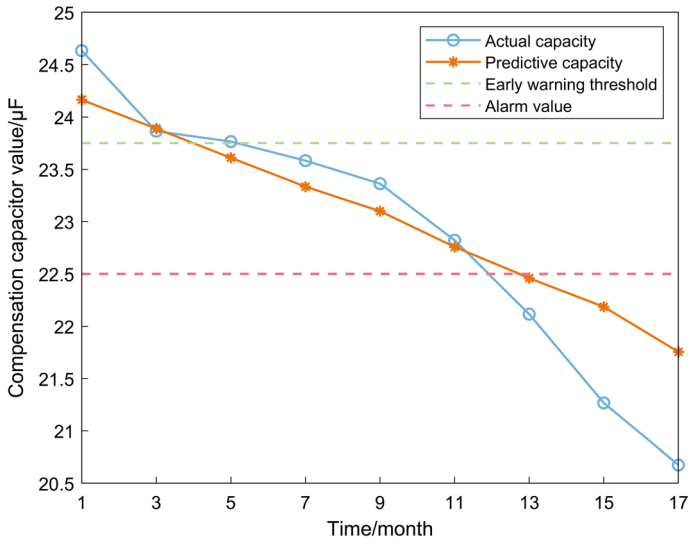
Fig. 13 Prediction results of compensation capacitor health state

thresholds and abnormal perception, and the fluctuation is within an acceptable range. Hence, the proposed failure warning mechanism is effective and feasible.

5 Conclusions

This paper proposes a health state prediction method based on deep learning, completing the compensation capacitor fault forecast applicable to complex and dynamic scenes of track circuits. The innovative work and conclusions are as follows:

(1) The degradation mechanism of compensation capacitors has been explored from long-term dependence, heterogeneity, and uncertainty knowledge. We establish a transmission state model for track circuits including compensation capacitors, and then introduce a defined difference function and a constructed degradation model to describe the failure law of capacitors.

(2) The methodology of SLCBN modeling is established. Focusing on the spatiotemporal characteristics of data, we optimize CNN, BiLSTM, and SSA to construct and integrate the SLCBN network framework based on regularization. Moreover, training and verification are completed on the compensation capacitor data set with invalidation mechanisms to achieve health state prognosis and mining. The results show that the SLCBN effect meets the best performance: RMSE $= 0.00705$, MSE $= 0.0005$, MAE $= 0.00504$, and MSE $= 0.02608$.

(3) The health state forecast can reflect the possibility of degradation and evolution of compensation capacitors. The constructed SLCBN provides a substantial body of knowledge for elucidating abnormal states and invalidation mechanisms. Through a series of calculations, the quantitative level of hazard threshold is

obtained, which can perceive the possibility of abnormal states and achieve a 100% accuracy rate for early warning.

In short, the deep neural network model built in this paper is helpful to effectively and intelligently extract numerous sequence features from historical monitoring information, successfully predicting the healthy state of compensation capacitors while improving maintenance efficiency, expected to benefit the "conditional repair" in intelligent maintenance of high-speed railway.

Author contribution CW contributed to the conceptualization, software, methodology, validation, writing—original draft, data curation, and resources. SY was involved in the conceptualization, methodology, supervision, writing—review and editing, and funding acquisition. CL assisted in the conceptualization, methodology, supervision, and writing—review and editing.

Data availability The authors do not have permission to share data.

Declarations

Competing interests The authors declare no competing interests.

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Ning B, Mo Z, Li K (2019) The intelligent technology application and development of the high-speed railway signal system. J Railw 41(03):1–9. https://doi.org/10.3969/j.issn.1001-8360.2019.03.001
2. Debiolles A, Oukhellou L, Aknin P (2004) Combined use of partial least squares regression and neural network for diagnosis tasks. In: 2004 Proceedings of the 17th International Conference on Pattern Recognition (ICPR), pp 573–576. https://doi.org/10.1109/ICPR.2004.1333837
3. Debiolles A, Denoeux T, Oukhellou L, Aknin P (2006) Output coding of spatially dependent sub-classifiers in evidential framework, Application to the diagnosis of railway track/vehicle transmission system. In: 2006 9th International Conference on Information Fusion, pp 1–6. https://doi.org/10.1109/ICIF.2006.301611
4. Oukhellou L, Aknin P, Delechelle E (2006) Railway infrastructure system diagnosis using empirical mode decomposition and hilbert transform. In: 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, pp 3615–3618. https://doi.org/10.1109/ICASSP.2006.1660866
5. Oukhellou L, Debiolles A, Denux T, Aknin P (2010) Fault diagnosis in railway track circuits using Dempster-Shafer classifier fusion. Eng Appl Artif Intell 23(1):117–128. https://doi.org/10.1016/j.engappai.2009.06.005
6. Xu K, Zhao L (2018) A rapid diagnosis method for multiple compensation capacitor faults of jointless track circuits. J Chin Railw Soc 40(02):67–72. https://doi.org/10.3969/j.issn.1001-8360.2018.02.010
7. Yang J, Zheng Q, Yao X, Wang X (2023) Multiple compensation capacitance fault location based on deep network transient characteristics of track circuit. J Railw Sci Eng. https://doi.org/10.19713/j.cnki.43-1423/u.T20221535

8. Feng D, Zhao L (2016) Method of estimation on capacitance of JTC compensation capacitor based on TCR monitoring data. J Chin Railw Soc 38(02):89–95. https://doi.org/10.3969/j.issn.1001-8360.2016.02.012

9. Wang S, Zhao L, Meng J, Wu N, Gao L (2021) Method of estimation on capacitance of JTC compensation capacitor based on data of inspection car. J Chin Railway Soc 43(12):55–61. https://doi.org/10.3969/j.issn.1001-8360.2021.12.007

10. Lyu G, Zhang H, Miao Q (2023) RUL prediction of lithium-ion battery in early-cycle stage based on similar sample fusion under lebesgue sampling framework. IEEE Trans Instrum Measur 72:1–11. https://doi.org/10.1109/TIM.2023.3260277

11. Hu L, Cg He, Cai Z, Wen L, Ren T (2019) Track circuit fault prediction method based on grey theory and expert system. J Vis Commun Image Represent 58:37–45. https://doi.org/10.1016/j.jvcir.2018.10.024

12. Simone D, Caputo E, Cinque M, Galli A, Moscato V, Russo S, Cesaro G, Criscuolo V, Giannini G (2023) LSTM-based failure prediction for railway rolling stock equipment. Expert Syst Appl 222:119767. https://doi.org/10.1016/j.eswa.2023.119767

13. Kang R, Wang JF, Chen JQ, Zhou JJ, Pang YZ, Gou LL, Cheng JF (2022) A method of online anomaly perception and failure prediction for high-speed automatic train protection system. Reliab Eng Syst Safe 226:108699. https://doi.org/10.1016/j.ress.2022.108699

14. Li T, Wu X, He Y, Peng X, Yang J, Ding R, He C (2023) Small samples noise prediction of train electric traction system fan based on a multiple regression-fuzzy neural network. Eng Appl Artif Intell 126:106781. https://doi.org/10.1016/j.engappai.2023.106781

15. Shi Y, Yang C, Wang J, Zhang Z, Meng F, Bai H (2023) A forecasting model of ionospheric fof2 using the LSTM network based on ICEEMDAN decomposition. IEEE Trans Geosci Rem Sens 61:1–16. https://doi.org/10.1109/TGRS.2023.3336934

16. Liao C, Chen C, Xiang C, Huang H, Xie H, Guo S (2022) Taxi-Passenger's destination prediction via GPS embedding and attention-based BiLSTM model. IEEE Trans Intell Transp Syst 23:4460–4473. https://doi.org/10.1109/TITS.2020.3044943

17. Rubasinghe O, Zhang X, Chau TK, Chow Y, Fernando T, Iu H (2023) A novel sequence to sequence data modelling based CNN-LSTM algorithm for three years ahead monthly peak load forecasting. IEEE Trans Pow Syst. https://doi.org/10.1109/TPWRS.2023.3271325

18. Zhu L, Shen C, Wang X, Liang H, Wang H, Tang T (2023) A learning based intelligent train regulation method with dynamic prediction for the metro passenger flow. IEEE Trans Intell Transp Syst 24:3935–3948. https://doi.org/10.1109/TITS.2022.3231838

19. Kavianpour P, Kavianpour M, Jahani E (2023) A CNN-BiLSTM model with attention mechanism for earthquake prediction. J Supercomput 79:19194–19226. https://doi.org/10.1007/s11227-023-05369-y

20. Chen J, Roberts C, Weston P (2008) Fault detection and diagnosis for railway track circuits using neuro-fuzzy systems. Control Eng Pract 16:585–596. https://doi.org/10.1016/j.conengprac.2007.06.007

21. Zhao L, Guo J, Li H, Liu W (2009) The simulation analysis of influence on jointless track circuit signal transmission from compensation capacitor based on transmission-line theory. In: 2009 3rd IEEE international symposium on microwave, antenna, propagation and EMC technologies for wireless communications, pp 1113–1118. https://doi.org/10.1109/MAPE.2009.5355926

22. Fu J, Wang X, Guo J (2018) Research on dynamic multi-wheel sets shunted track circuit modeling. J Railw Sci Eng 15(09):2374–2384. https://doi.org/10.19713/j.cnki.43−1423/u.2018.09.026

23. Xie Y, Yang S, Liu C, Wang C (2022) Ballast resistance estimation method based on one dimensional convolutional neural network. In: 2022 IEEE 11th Data Driven Control and Learning Systems Conference (DDCLS), pp 164–169. https://doi.org/10.1109/DDCLS55054.2022.9858574

24. Ministry of Railways of the People's Republic of China (2013) Cab signal systems on board: TB 3287–2013. China Railway Publishing House, Beijing

25. Ministry of Railways of the People's Republic of China (2008) Technical standard for railway signal maintenance rules. China Railway Press, Beijing, pp 92–99

26. Meng J, Hou Y (2018) Failure mode analysis of compensation capacitor of track circuit. Chin Railw 673(07):93–95. https://doi.org/10.19549/j.issn.1001-683x.2018.07.093

27. Li W (2017) High-speed railway train operation control technology—ZPW-2000 series of jointless track circuits, 1st edn. China Railway Press, Beijing

28. GB/T 25121.2-2018. Rail transportation locomotive equipment electronic capacitor part 2: non-solid electrolytic aluminum electrolytic capacitors

29. Qin Q, Zhao S, Chen S, Huang D (2018) Remaining useful life prediction of electrolytic capacitor based on particle swarm optimization particle filter. Comput Eng Appl 54(20):237–241. https://doi.org/10.3778/j.issn.1002-8331.1706-0386

30. Liu C, Yang S (2022) Using text mining to establish knowledge graph from accident/incident reports in risk assessment. Expert Syst Appl 207:117991. https://doi.org/10.1016/j.eswa.2022.117991

31. He Y, Zhang L, Chen Z (2023) A framework of structural damage detection for civil structures using a combined multi-scale convolutional neural network and echo state network. Eng Comput 39:1771–1789. https://doi.org/10.1007/s00366-021-01584-4

32. Liu C, Yang S, Chu S, Wang C, Liu R (2022) Application of ensemble learning and expert decision in fuzzy risk assessment of railway signaling safety. In: 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), pp 3691–3697, https://doi.org/10.1109/ITSC55140.2022.9922049

33. Feng J, Liu L, Hou X, Pei Q, Wu C (2023) QoE fairness resource allocation in digital twin-enabled wireless virtual reality systems. IEEE J Sel Areas Commun 41:3355–3368. https://doi.org/10.1109/JSAC.2023.3313195

34. Yue YG, Cao L, Lu DW, Hu ZY, Xu MH, Wang SX, Li B, Ding HH (2023) Review and empirical analysis of sparrow search algorithm. Artif Intell Rev 56:1–53. https://doi.org/10.1007/s10462-023-10435-1

35. Zhao ZN, Yun SN, Jia LY, Guo JX, Meng Y, He N, Li XJ, Shi J, Yang L (2023) Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features. Eng Appl Artif Intell 121:105982. https://doi.org/10.1016/j.engappai.2023.105982

Authors and Affiliations

Conghui Wang[1] · Shiwu Yang[1] · Chang Liu[2]

✉ Shiwu Yang
ysw@bjtu.edu.cn

Conghui Wang
21111071@bjtu.edu.cn

Chang Liu
19111061@bjtu.edu.cn

[1] School of Automation and Intelligence, Beijing Jiaotong University, Beijing 100044, China

[2] CRSC Research & Design Institute Group Co., Ltd, Beijing 100070, China