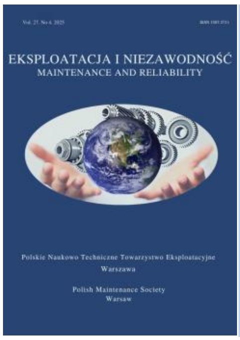

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 27 (2025), Issue 4

journal homepage: http://www.ein.org.pl

Article citation info:
Ma D, Wang Z, He T, Electromagnetic Relay Contact Resistance Prediction Based on TimeGAN with CNN-LSTM-Attention, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2025: 27(4) http://doi.org/10.17531/ein/204610

Electromagnetic Relay Contact Resistance Prediction Based on TimeGAN with CNN-LSTM-Attention

Indexed by: Web of Science Group

Dongkun Ma[a], Zhaobin Wang[a,b,*], Tianyang He[a]

[a] College of Automation Engineering, Jiangsu University of Science and Technology, China
[b] School of Electronic and Electrical Engineering, University of Leeds, United Kingdom

Highlights

- Proposes TimeGAN-CNN-LSTM-Attention model for precise contact resistance prediction.
- TimeGAN-generated synthetic data validated via PCA/KDE to resolve degradation data.
- Combines adversarial data generation and deep learning in resource-limited systems.
- Achieves 30.22% R² boost , demonstrating superior prediction accuracy.

Abstract

Electromagnetic relays are critical in aerospace and military systems, affecting the safety and stability of applications like aircraft control, satellite communication, and missile launchers. However, the scarcity of degradation data and complex variations in contact resistance pose challenges. Traditional methods often struggle with small samples. To address these issues, propose a novel framework integrating TimeGAN with a CNN-LSTM-Attention model. TimeGAN generates synthetic degradation data that aligns with the statistical distribution of the original dataset, mitigating data scarcity. Data quality is evaluated using PCA and KDE. The CNN-LSTM model captures multi-scale temporal features, while the attention mechanism highlights critical features to improve contact resistance prediction accuracy. Experimental results show that the proposed framework outperforms traditional methods, demonstrating robust performance even without data augmentation. These findings offer a valuable foundation for health monitoring and fault prediction in high-reliability systems.

Keywords
electromagnetic relay, contact resistance, TimeGAN, CNN-LSTM, attention mechanism.

1. Introduction

Electromagnetic relays play a pivotal role in aerospace and military applications. In the aerospace sector, they are widely employed in satellite communication, attitude control, and the operation of propulsion systems, where their high reliability ensures the long-term stability of spacecraft. In military applications, electromagnetic relays are critical components in missile launching systems and radar equipment, facilitating high-precision signal switching and control—both essential for the proper functioning of key systems. These relays are designed for extended storage durations and exhibit slow degradation, making it challenging to accumulate sufficient failure data within a short period due to the gradual degradation process and the influence of multiple external factors. Traditional reliability assessment methods often struggle to accurately evaluate their long-term storage reliability[1-2]. Although accelerated degradation testing has emerged as an important research approach, uncertainties in degradation mechanisms pose significant challenges in selecting appropriate stress conditions and developing suitable test equipment. Furthermore, issues such as high testing costs, prolonged testing

(*) Corresponding author.
E-mail addresses: D. Ma (ORCID: 0009000453184656) 13315964094@163.com, Z. Wang (ORCID: 0000000237678179) wangzb@just.edu.cn, T. He (ORCID:0009000708490699) 17195391@cumt.edu.cn,

cycles, and limited sample availability further complicate research efforts[3-4].

Contact resistance serves as a critical indicator of electromagnetic relay degradation, with its temporal evolution directly dictating operational reliability and lifespan. During repeated switching operations, mechanical wear and electrical arcing induce surface roughening and oxide layer formation at the contact interface[5]. This progressive deterioration elevates contact resistance, leading to increased Joule heating and localized temperature rise, which further accelerates material oxidation—a self-reinforcing degradation mechanism[6]. Experimental studies demonstrate that a 20% increase in contact resistance correlates with a 35% reduction in relay lifespan under rated loads, as excessive voltage drops ($\Delta V > 50$ mV) compromise signal integrity in aerospace control systems. Traditional reliability models often oversimplify these nonlinear interdependencies, treating resistance as a static parameter rather than a dynamic precursor to failure. This gap motivates our data-driven approach to capture the spatiotemporal degradation signatures encoded in resistance trajectories, enabling proactive health management.

Data augmentation techniques have been widely applied in regression tasks and image classification, yet their adoption for time-series data remains relatively underexplored[7-8]. Generative Adversarial Networks (GANs)[9] offer a promising approach to data augmentation by generating synthetic samples that align with the statistical distribution of real data through an adversarial learning process. Building upon this foundation, Yoon et al. proposed Time-series Generative Adversarial Networks (TimeGAN)[10], which combine supervised and unsupervised learning techniques to model temporal dependencies while maintaining the diversity of distributions. Prior research has demonstrated the effectiveness of TimeGAN in addressing data scarcity challenges; for instance, one study[11] employed TimeGAN to generate synthetic fault samples resembling historical failure data, mitigating the issue of limited failure samples. In high-reliability applications, another study[12] combined TimeGAN with a CNN-BiLSTM-Attention model to generate health indicators for electro-hydrostatic actuators (EHAs). Additionally, research[13] introduced Extraction TimeGAN to enhance aero-engine failure data, validating its performance using Principal Component Analysis (PCA) to ensure consistency with real data distributions. Furthermore, a Conditional Time-Series Generation Adversarial Network (C-TimeGAN)[14] was proposed to generate high-quality equipment trajectories by imposing specific constraints, thereby improving the accuracy of Non-Intrusive Load Monitoring (NILM) models.

In this paper, we propose a contact resistance prediction method for electromagnetic relays based on TimeGAN and a CNN-LSTM-Attention framework to address the challenge of insufficient training data due to the scarcity of degradation samples. First, TimeGAN is utilized to generate synthetic degradation data that maintain statistical consistency with the original dataset, thereby expanding the training set. The distribution consistency of the generated data is verified using Principal Component Analysis (PCA) and Kernel Density Estimation (KDE). Second, a predictive model combining CNN, LSTM, and Attention mechanisms is constructed to extract multi-scale spatiotemporal features and key time-step information from the degradation data, enhancing the modeling capability for complex degradation processes. Finally, both the generated and original data samples are used to train the proposed model, and its performance is compared with the traditional CNN-LSTM model, standard LSTM model, GRU model and Transformer model. This approach establishes a novel paradigm for prognostics and health management (PHM) of electromechanical systems under data-limited conditions, effectively bridging adversarial data generation with deep feature learning to advance reliability prediction in resource-constrained industrial applications.

2. Related Mechanism Analysis

2.1. TimeGAN

The proposed TimeGAN framework demonstrates an innovative integration of unsupervised adversarial generation training (characteristic of GAN architectures) with supervised regression-based learning to effectively enhance datasets for analyzing degraded parameters of electromagnetic relays. The methodology systematically categorizes degradation data features into two distinct types: static features $\boldsymbol{S}$ and dynamic features $\boldsymbol{X}$, which respectively occupy orthogonal vector spaces ζ and χ.Static features refer to time-invariant parameters that remain stable throughout the relay's lifecycle, such as material properties , geometric configurations, and initial operational states. These features characterize the intrinsic attributes of the electromagnetic relay. In contrast, dynamic features describe time-evolving parameters influenced by operational wear and environmental interactions, including transient contact resistance fluctuations, pickup/release times, over-travel displacement, and temperature-dependent degradation trends. The potential feature distributions $\hat{p}(\boldsymbol{S}, \boldsymbol{X}_{1:T})$ are learned from the real degradation data so that they are as close as possible to the actual distributions $p(\boldsymbol{S}, \boldsymbol{X}_{1:T})$ of the static and dynamic feature vectors.

The TimeGAN architecture combines an autoencoder framework with adversarial training mechanisms to synthesize high-fidelity temporal data. The autoencoder component comprises an embedding network that compresses relay degradation data into a low-dimensional latent space and a recovery network that reconstructs the original data space, ensuring preservation of critical degradation features. Simultaneously, the adversarial network employs a generator that synthesizes realistic time-series data from noise-conditioned latent vectors, while a discriminator evaluates the statistical congruence between generated and empirical distributions [15]. This dual mechanism jointly optimizes feature reconstruction fidelity and distributional alignment, achieving a robust synthesis of degradation patterns as detailed in figure 1.

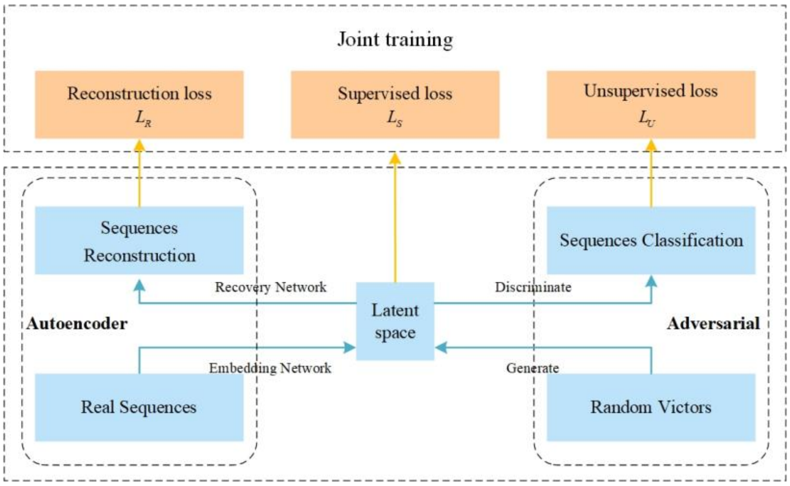

Figure 1. Structure of TimeGAN.

TimeGAN effectively models the temporal dynamics of contact resistance degradation through a composite loss architecture integrating three critical components. The reconstruction loss (L_R) preserves fundamental distributional characteristics of empirical resistance data by minimizing autoencoding errors. A physics-guided supervision loss (L_S) enforces consistency between generated sequences and domain-specific degradation models, ensuring adherence to known electromechanical principles. Concurrently, the adversarial loss (L_U) facilitates cross-parameter correlation learning, capturing dynamic interactions between resistance evolution and auxiliary degradation indicators—including suction/release times and over-travel displacement—through discriminator-guided feedback mechanisms. By jointly optimizing these complementary objectives, the framework synthesizes temporally coherent degradation trajectories that maintain both local feature fidelity and global physical plausibility, as demonstrated through comparative analysis of synthetic and empirical spatiotemporal patterns. The three loss functions are shown below:

$$L_R = E_{\mathbf{S},\mathbf{X}_{1:T \sim P}}\left[\left\|\mathbf{S}-\tilde{\mathbf{S}}\right\|_2 + \sum t\left\|\mathbf{X}_t-\widetilde{\mathbf{X}_t}\right\|_2\right] \tag{1}$$

$$L_s = E_{\mathbf{S},\mathbf{X}_{1:T \sim P}}\left[\left\|\mathbf{h}_t - g_\chi(\mathbf{h}_s, \mathbf{h}_{t-1}, z_t)\right\|_2\right] \tag{2}$$

$$L_U = E_{\mathbf{S},\mathbf{X}_{1:T \sim P}}[\log y_s + \textstyle\sum_t \log y_t] + E_{\mathbf{S},\mathbf{X}_{1:T \sim P}}[\log(1 - \widehat{y_s}) + \textstyle\sum_t \log(1-\widehat{y_t})] \tag{3}$$

2.2. CNN-LSTM model

The model integrates Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM). The CNN layer uses a multi-scale convolutional kernel to extract local morphological features, which effectively identifies precursors to sudden changes in contact resistance. The two-layer LSTM structure captures long-term dependencies through its memory cell updating mechanism, enabling accurate localization of the transition point from the stable to the accelerating period. Zhi et al.[16]constructed a harmonic reducer fault detection model based on the CNN-LSTM framework and a novel noise reduction algorithm (WRCTD), achieving noise reduction, feature extraction, and accurate fault classification for complex industrial vibration signals.Xu et al.[17] introduced a hybrid model based on EMD-CNN-LSTM, aimed at improving the accuracy of short-term power load forecasting, particularly for the nonlinear and nonsmooth characteristics of load sequences. In this study, the CNN module consists of two convolutional layers and two pooling layers that alternate with each other. This design combines the principles of regional connectivity and weight sharing. The convolutional layer extracts latent features from the degraded data, with the CNN's convolution kernel being one-dimensional. This design is suitable since the contact resistance, pickup time, release time, and overtravel time of the electromagnetic relay are represented as one-dimensional time series.The pooling layer serves to compress the high-dimensional feature maps generated during feature extraction, effectively compressing spatial dimensions while preserving critical data attributes. This dimensionality reduction mechanism lowers computational costs in subsequent processing stages and enhances model generalization by suppressing redundant parameters that may contribute to overfitting. The structure of the CNN-LSTM model is illustrated in figure 2.

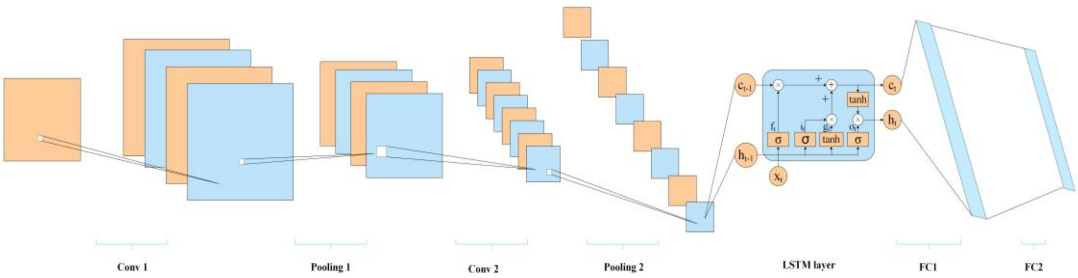

Figure 2. Structure of CNN-LSTM model.

LSTM, as a variant of RNN, can efficiently capture dynamic changes and long-term dependencies in degraded data through its unique gating mechanism, which consists of three parts: input gate, forgetting gate, and output gate. Its calculation formula is as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{4}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{5}$$

$$\bar{C}_t = tanh(W_c[h_{t-1}, x_t] + b_c) \tag{6}$$

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = O_t\, tanh(C_t) \tag{8}$$

In the temporal gating architecture, input x_t and hidden state h_t are processed through three regulatory components: forget gate f_t, input gate i_t ,and output gate O_t. Each gate operates through distinct parameter sets - weight matrices W_f, W_i, W_c, W_o and corresponding bias vectors b_f, b_i, b_c, b_o govern information flow regulation. The compressed feature maps from the pooling layer feed into a 128-unit LSTM network, followed by dimensionality transformation through a dense output layer for prediction generation. To optimize model robustness, dropout regularization is strategically implemented between network layers, effectively suppressing co-adaptation of hidden units. The training protocol employs two adaptive mechanisms:1) Early termination based on validation plateau detection.2) Cosine-annealed learning rate scheduling, collectively enhancing convergence efficiency while maintaining solution stability.

2.3. Attention mechanism

The Attention mechanism, inspired by biological selective perception principles, dynamically allocates feature weights to enable models to focus on critical information. This approach demonstrates cross-domain applicability in reliability engineering: In typhoon impact prediction, attention-enhanced CNN architectures process multimodal geospatial data (meteorological patterns, topographic features) combined with XGBoost to improve robustness in power grid disaster warnings[18]. For lithium battery health monitoring, an attention-guided CNN-LSTM framework addresses insufficient temporal feature extraction in state-of-health (SOH) estimation through adaptive temporal saliency mapping[19]. Applied to contact resistance prediction, the mechanism automatically identifies critical degradation phases while suppressing noise interference via learnable weight allocation, enhancing temporal pattern recognition without manual feature engineering. Cross-domain applications validate its effectiveness in extracting physically interpretable features from complex spatiotemporal degradation processes[20]. The calculation formula of the attention mechanism is:

$$\text{score}(h,\bar{h}_s) = \bar{h}W\bar{h}_s \tag{9}$$

$$\alpha_{ls} = \frac{exp(\text{score}(h,\bar{h}_s))}{\sum_{s=1}^{S\sum \text{scor}\bar{\text{e}}_s} exp} \tag{10}$$

$$C_t = \textstyle\sum_s \alpha_{ls}\,\bar{h}_s \tag{11}$$

$$a_t = f(c_t, h_t) = tanh(W_c[c_t; h_t]) \tag{12}$$

Where score$(h,\bar{h}_s)$ is the weight score, α_{ls} is the weight, $\bar{h}_s$ is the hidden variables of all original moments, C_t denotes the Attention-weighted feature vector, a_t is the attention vector, and h_t is the hidden variables of the current moment. In this paper, the Attention mechanism is integrated into the dual LSTM layer, and the unit layer structure is schematically shown in figure 3.

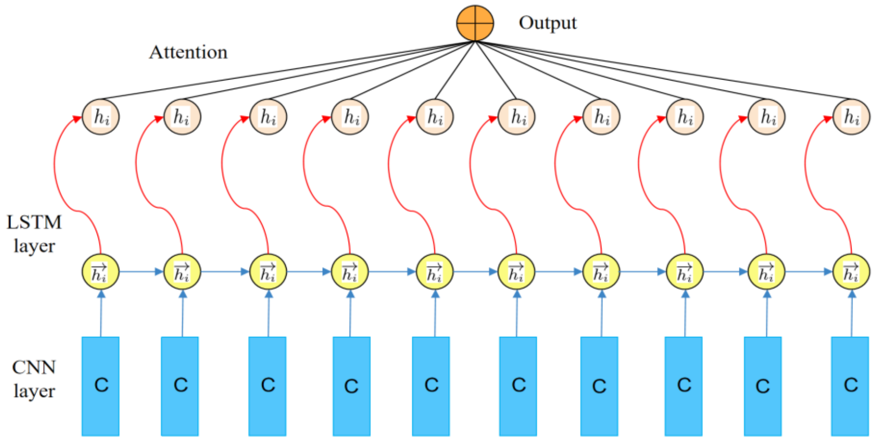

Figure 3. Schematic diagram of CNN-LSTM-Attention unit layer structure.

3. TimeGAN-CNN-LSTM-Attention Model Construction

In this paper, a total of 171 sets of electromagnetic relay degradation data are used, and less training data will cause overfitting problem of the deep learning model, in this regard, a contact resistance prediction method of electromagnetic relay based on TimeGAN and CNN-LSTM-Attention model is proposed, which utilizes TimeGAN for data enhancement of original degradation data, and at the same time extracts global features of the degradation data through CNN-LSTM-Attention to capture important time steps, to accurately achieve the contact resistance prediction. CNN-LSTM to extract the global features of the degraded data, combined with the Attention mechanism to assign weights to capture the important time steps, so as to accurately realize the prediction of contact resistance. The specific prediction process of TimeGAN-CNN-LSTM-Attention is shown as follows:

(1) The degraded data is fed into the TimeGAN model for training, where unsupervised adversarial generation from GAN is combined with supervised training in the regression model. The difference between the generated data and the original degraded data is evaluated in three dimensions using the constructed loss functions, and the process iterates until the difference meets the specified criteria.

(2) Normalize the original degraded data and the new data generated via TimeGAN.

(3) The normalized dataset is fed into the CNN-LSTM hybrid architecture through a systematic computational pipeline. The convolutional stage employs sequential operations with 3×3 kernel filters across alternating layers, generating 64-channel feature maps that undergo spatial dimension compression via 2×2 pooling operations. These hierarchically abstracted representations are subsequently propagated into a two-tiered LSTM module: the initial recurrent layer with 128 memory units integrates an attention mechanism to dynamically weight temporal dependencies, while the subsequent LSTM layer maintains identical 128-unit capacity to preserve sequential pattern integrity. The processed temporal features are then projected through a fully connected regressor to estimate contact resistance values. Throughout this computational

cascade, rectified linear unit (ReLU) activations govern nonlinear transformations across all learnable components. Upon completing the forward propagation setup, the Adam optimizer orchestrates parameter updates during backpropagation to minimize prediction errors, adhering to its inherent adaptive learning rate properties for stable convergence.

(4) The 1000 sets of newly degraded data generated by the TimeGAN model in (1) are input into the constructed CNN-LSTM-Attention model in (3) as a training set, and the test set is extracted from the original data. Eventually, the data are back-normalized to accurately achieve the prediction of the contact resistance of the electromagnetic relay.

To ensure the reproducibility of the experiments, the hyperparameter configurations for the CNN-LSTM-Attention model architecture and training protocol are detailed in table 1:

Table 1. Model implementation details.

Component	Parameter	Value/Range
CNN Module	Number of Convolutional Layers	2
	Kernel Size	3×3
	Number of Kernels	64 per layer
	Activation Function	ReLU
	Pooling Type	Max Pooling
	Pooling Window	2×2
LSTM Module	Number of Recurrent Layers	2
	Hidden Units per Layer	128
	Dropout Rate	0.2
Attention Mechanism	Attention Type	Additive Attention
	Weight Initialization	Xavier Normal
	Context Vector Dimension	64
Training Strategy	Optimizer	Adam
	Initial Learning Rate	0.001
	Batch Size	24
	Training Epochs	100
	Learning Rate Schedule	Cosine Annealing

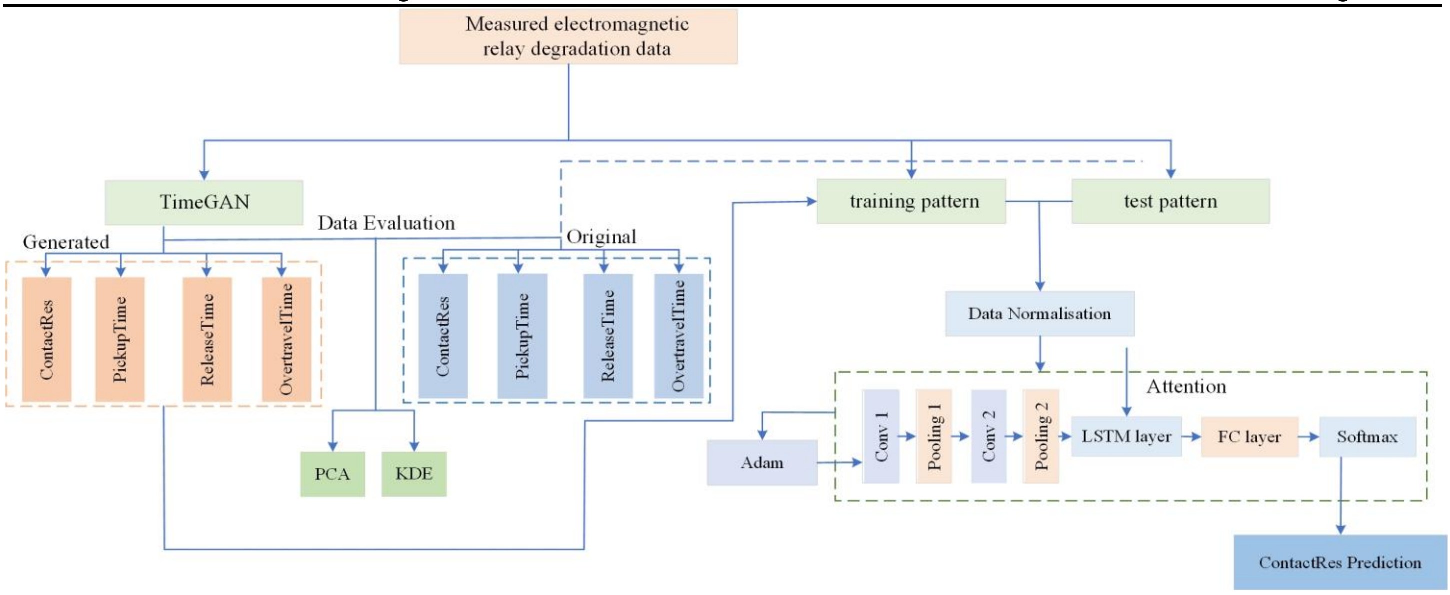

Figure 4. Flowchart for TimeGAN-CNN-LSTM-Attention model construction.

The proposed CNN-LSTM-Attention model is trained on an NVIDIA Tesla V100 GPU with 32GB VRAM, utilizing PyTorch 2.0.1 and CUDA 11.7. The training process require approximately 1.5 hours for 100 epochs with a batch size of 24, achieving convergence at epoch 78 (early stopping triggered). The model contains 500 thousand parameters, with the LSTM module accounting for 68%, the CNN module 22%, and the attention mechanism 10%. For comparative analysis, the baseline LSTM and CNN-LSTM models require 1.1 hours and 1.2 hours, respectively, under identical hardware configurations. While the parameter scale exceeds lightweight architectures, the computational overhead remains manageable for cloud-based deployment. For edge devices with limited resources, future work may explore model pruning or quantization to reduce memory footprint without significant accuracy loss.

The specific TimeGAN-CNN-LSTM-Attention model construction flowchart is shown in figure 4.

4. Experimental results and analysis

4.1. Primary degradation data collection

The experimental data was collected from a specific model of electromagnetic relay tested on the accelerated degradation test platform at Harbin Institute of Technology. This testing period spanned from January 18, 2012, to January 28, 2013, with degradation data recorded every 48 hours. The data includes measurements of contact resistance, pickup time, release time, and over-travel time. In this paper, contact resistance is treated as the primary target for prediction, while the other three metrics serve as multi-step inputs for the model. During the process of predicting contact resistance, it was noted that differences in units and magnitudes could impact prediction accuracy. To enhance the model's convergence and stability, all data were normalized to a range of [0, 1]. After normalization, the data will need to be back-normalized using the following formula:

$$x' = \frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{13}$$

$$x = x' \times (x_{\max} - x_{\min}) + x_{\min} \tag{14}$$

Where x, x' are the data before and after normalization, $x_{\max}$,$x_{\min}$ are the maximum and minimum values in the data before normalization.

4.2. Quality assessment of generated data

To better evaluate the quality of the 1000 newly generated data sets and their suitability for practical use, the expanded data is analyzed using Principal Component Analysis (PCA) and Kernel Density Estimation (KDE). PCA is employed to reduce the dimensionality of the data, focusing on eliminating degraded components that could affect the reliability of electromagnetic relay storage. This reduction allows for a comparison of the feature distribution between the generated and original datasets in a two-dimensional space. Figure 5 presents the 2D feature distribution of both the generated and original data after dimensionality reduction, showing significant overlap between the two. KDE is utilized to examine the relative data density across different intervals. By overlaying multiple datasets, it becomes easier to assess the distributional differences and understand the gap between the model-generated data and the original data. Figure 6 illustrates that the KDE fitting is effective and aligns well with the original data.

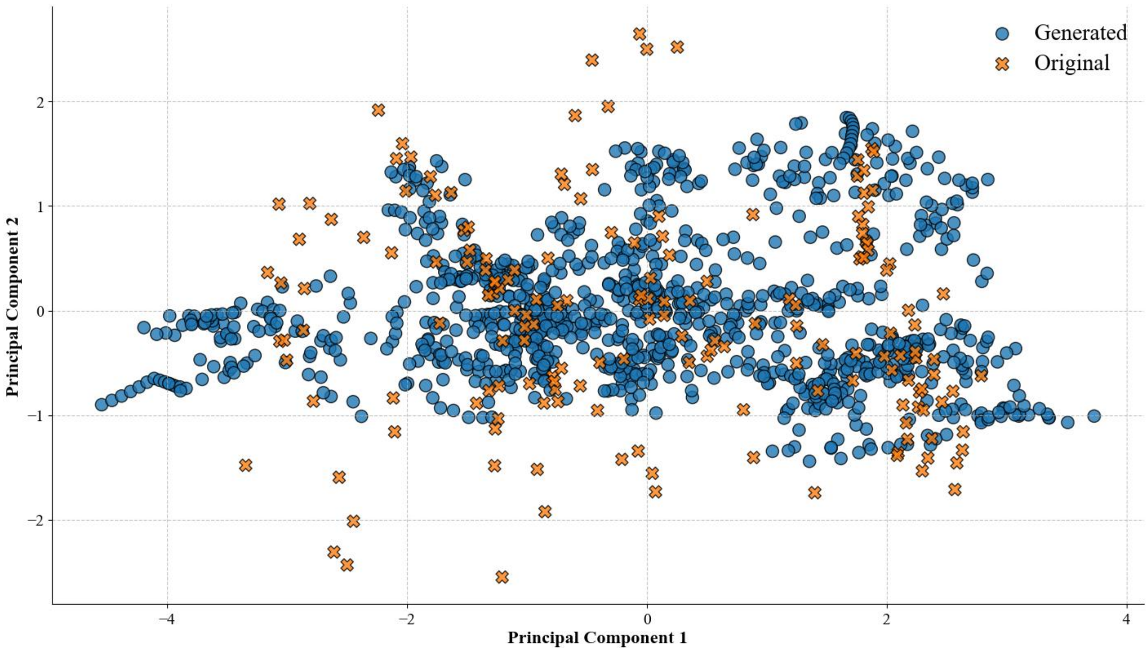

Figure 5. Principal Component Analysis plot.

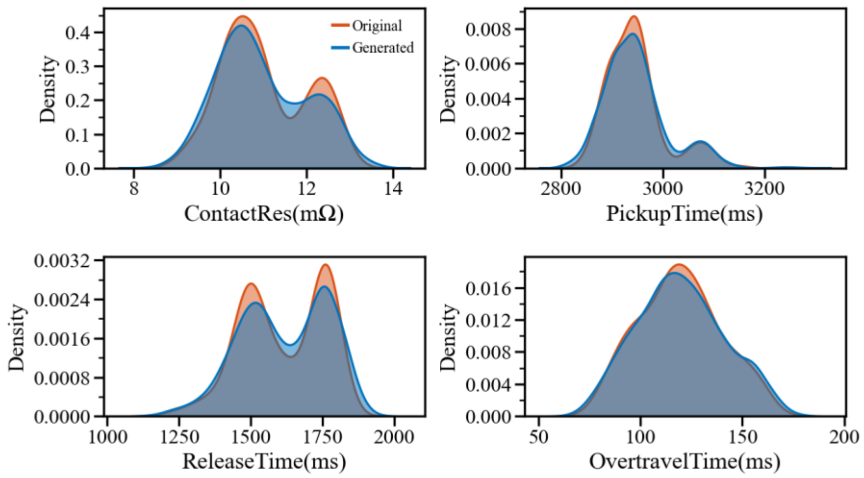

Figure 6. Kernel Density Estimation plot.

4.3. Attention Weight Analysis

In the context of electromagnetic relay degradation, critical time steps refer to specific operational intervals where abrupt changes in contact resistance or auxiliary parameters signal accelerated wear or impending failure. These steps are characterized by transient phenomena such as micro-arcing, oxide layer breakdown, or contact surface roughening, which precede macroscopic performance deviations. The proposed CNN-LSTM-Attention model explicitly identifies these steps through its attention mechanism, dynamically assigning higher weights to temporal features indicative of degradation acceleration. To validate the interpretability of the attention mechanism in identifying critical degradation phases of 1000 groups, we visualize the temporal attention weights of the test set samples and correlate them with domain-specific degradation events in figure 7.

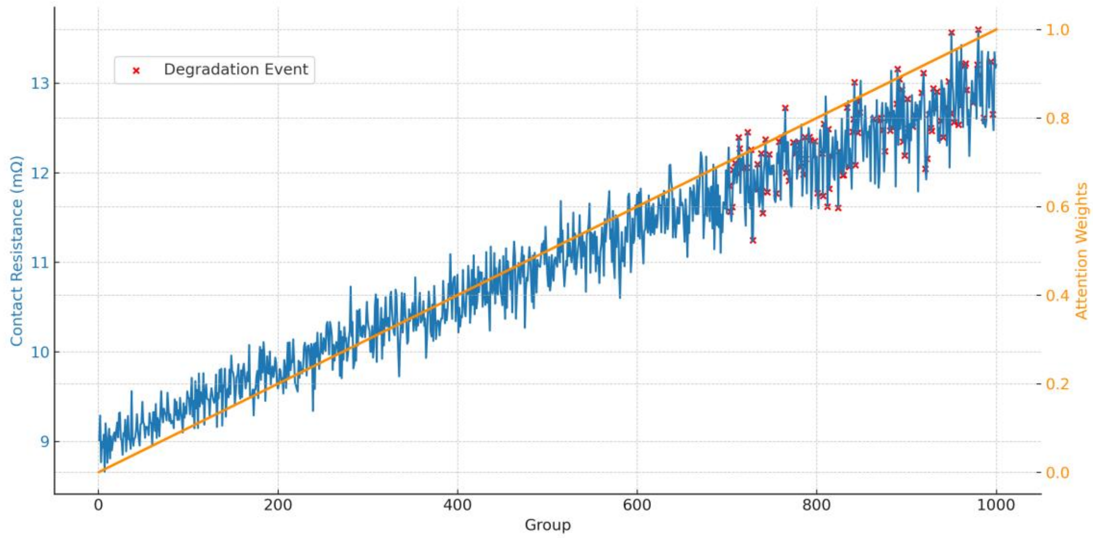

Figure 7. Temporal attention weights correlation with degradation phases.

The red markers represent degradation events, and as shown, these events coincide with high attention weight values, particularly in the latter time steps. This correlation suggests that the model assigns more importance to time periods that are crucial for the degradation process, aligning with our domain knowledge of how degradation accumulates over time, particularly under operational stress conditions. This visual analysis helps clarify the relationship between attention weights and degradation events, making it evident that the model indeed focuses on critical time steps where degradation processes are most prominent, thus validating the model's ability to capture these important patterns.

4.4. Evaluation indicators

Commonly used metrics in the evaluation of prediction algorithms include goodness of fit (R^2), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). These indices collectively quantify deviations between model outputs and ground truth observations, systematically assessing predictive performance and error distribution characteristics.

R^2, also known as the coefficient of determination, is a statistical measure that quantifies how well a regression model explains the variability in the data. It ranges from 0 to 1, with higher values indicating a better fit and greater explanatory power of the model. A value closer to 1 suggests that the model accounts for a large proportion of the variance in the observed data. The formula for R^2 is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\overline{y}_i - y_i)^2} \tag{15}$$

MAE is a metric that calculates the average of the absolute differences between the predicted values and the actual values, providing a measure of the average size of prediction errors. A lower MAE indicates more accurate predictions. The formula for MAE is as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{16}$$

MAPE is a metric that measures the relative error between predicted and actual values. It calculates the absolute difference between the predicted and actual values, expresses this error as a percentage, and then averages the percentages. A lower MAPE indicates better prediction accuracy. The formula for MAPE is as follows:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{17}$$

RMSE is the square root of the Mean Squared Error (MSE) and offers a measure of the error magnitude that aligns with the scale of the original data. It provides a direct indication of how far the predictions deviate from the actual values. The formula for RMSE is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{18}$$

In the formulas $\hat{y}_i$，y_i，$\overline{y}_i$ are the model predictions, the original values, and the average of the original values, respectively.

4.5. Analysis of results

Prior to model construction, a dataset consisting of 171 groups of degraded electromagnetic relay data is organized. The first 124 groups are designated as the training set, while the remaining 47 groups are used for testing. To evaluate the performance of the CNN-LSTM-Attention model, it is compared with the traditional CNN-LSTM model, standard LSTM model, GRU model and Transformer model with the goal of achieving more accurate contact resistance predictions. Furthermore, to assess the impact of TimeGAN on data augmentation, 1000 new sets of degraded data are generated by TimeGAN, based on the original 171 groups. These newly generated datasets are then incorporated into the training set, while the original 47 groups remained as the test set. The data is input into the CNN-LSTM-Attention model for contact resistance prediction. During training, a batch size of 24 samples is used for each weight update, and the training duration is set to 100 epochs to avoid both underfitting and overfitting.The comparison of the different contact resistance prediction methods, both before and after data enhancement by TimeGAN，is illustrated in figure 8 and 9.

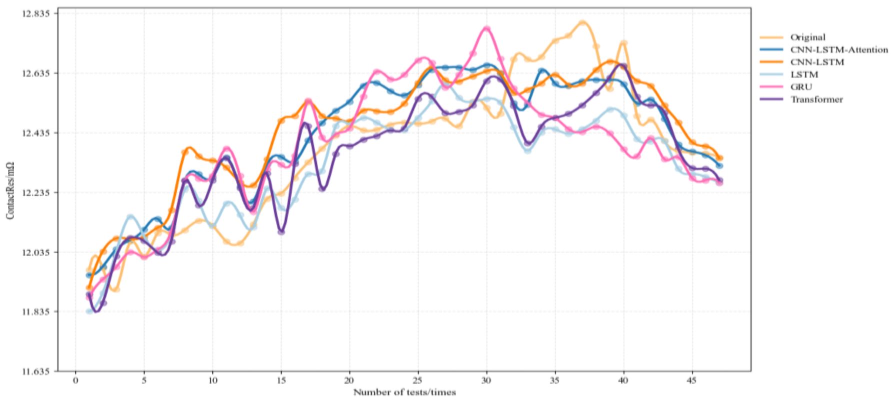

Figure 8. Comparison of contact resistance prediction of different models before enhancement.

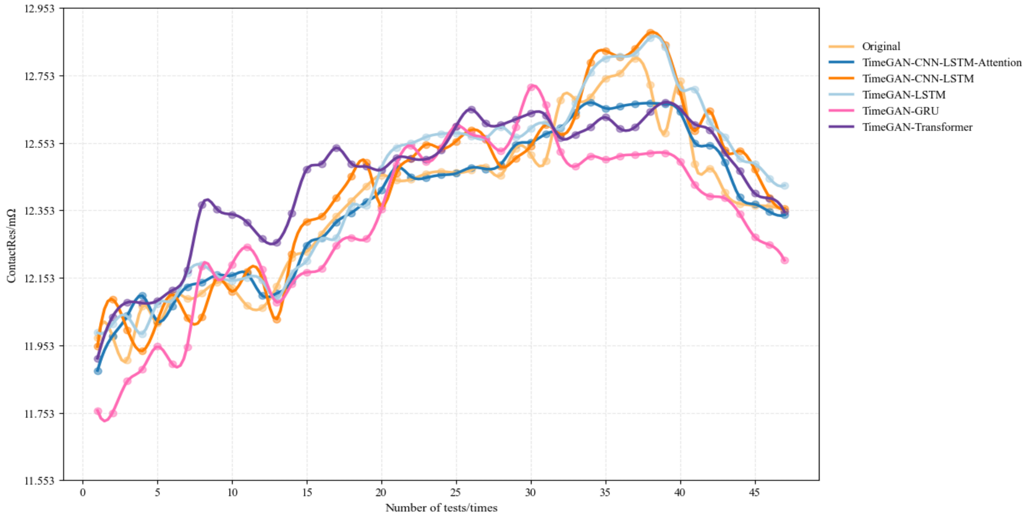

Figure 9. Comparison of contact resistance prediction of different models after enhancement.

TimeGAN-LSTM, TimeGAN-CNN-LSTM, TimeGAN-GRU, TimeGAN-Transformer,TimeGAN-CNN-LSTM-Attention and the specific evaluation metrics of each model before enhancement are shown in table 2.

Table 2. Specific evaluation indicators for the model.

Model	R^2	MAE	MAPE	RMSE
LSTM	0.6741	0.0971	0.7773	0.1346
CNN-LSTM	0.6782	0.1158	0.9378	0.1338
GRU	0.4877	0.1415	1.1454	0.1688
Transformer	0.7147	0.0994	0.0159	0.1259
CNN-LSTM-Attention	0.7322	0.1047	0.8445	0.1220
TimeGAN-LSTM	0.8287	0.0807	0.6513	0.0976
TimeGAN -CNN-LSTM	0.8522	0.0751	0.6055	0.0907
TimeGAN-GRU	0.6043	0.1258	0.0220	0.1483
TimeGAN- Transformer	0.7425	0.1013	0.8171	0.1197
TimeGAN-CNN-LSTM-Attention	0.9534	0.0396	0.3194	0.0509

In terms of the performance of the base model, the gradual optimisation of LSTM, CNN-LSTM and CNN-LSTM-Attention improves the R^2 from 0.6741 to 0.7322, and the RMSE decreases by 9.3%, which verifies the effectiveness of the convolutional layer with the attention mechanism for temporal feature extraction. However, the original GRU model performs weakly (R^2=0.4877), while Transformer shows strong competitiveness (R^2=0.7147), indicating that the self-attention mechanism has a unique advantage in the unenhanced data. Enhanced by the introduction of TimeGAN data, the performance of all models jumped significantly, with TimeGAN-GRU (R^2=0.6043) and TimeGAN-Transformer (R^2=0.7425) showing a limited increase, though improved from the base version. While TimeGAN-CNN-LSTM-Attention stands out with an overwhelming advantage, with R^2 reaching 0.9534, a 30.2% improvement over the pre-enhancement version, RMSE sharply decreasing by 58.3% to 0.0509, and MAE (0.0396) and MAPE (31.94%) decreasing by more than 60%, with breakthroughs leading in all four metrics. This proves that the deep synergy between temporal generation adversarial network and CNN-LSTM-Attention architecture can not only expand the diversity of temporal patterns through data enhancement, but also accurately capture the deep temporal dependency through multi-scale feature extraction and dynamic weight allocation mechanism, and ultimately achieve the double qualitative change of prediction accuracy and stability.

5. Conclusions

1) Propose a prediction method based on TimeGAN with CNN-LSTM-Attention. The high-quality time-series data generated by TimeGAN substantially expands the training dataset, effectively tackling the issue of small sample sizes. Experimental results demonstrate that the model's goodness-of-fit (R^2) increases from 0.7322 to 0.9534, while the root mean square error (RMSE) drops from 0.1220 to 0.0509 after incorporating TimeGAN. These enhancements emphasize the critical role of data augmentation in accurately modeling complex degradation features.

2) The constructed CNN-LSTM-Attention model integrates the feature extraction capabilities of convolutional neural networks, the time series dependency modeling abilities of LSTMs, and the attention mechanism's capability to focus on key time steps. This combination effectively enhances the modeling accuracy of contact resistance time series data.

Compared to the traditional LSTM model, the method presented in this study improves the R^2 value by 41.45% and reduces the RMSE by 62.17%. These enhancements significantly decrease prediction errors and highlight the model's potential for application in the health monitoring of high-reliability equipment.

3) Comparative analysis before and after enhancement shows that the method in this study has significant advantages in data-scarce scenarios. The mean absolute error (MAE) of the TimeGAN-CNN-LSTM-Attention model decreases from 0.1047 to 0.0396, and the relative percentage error (MAPE) decreases from 0.8445 to 0.3194, with an error reduction of more than 60%, verifying the feasibility and effectiveness of data augmentation combined with deep learning models.

Although this study achieved good results in model accuracy and small sample scenarios, there are still some limitations: (1) the quality of TimeGAN-generated data is greatly affected by the distribution of the original data and noise, which may reduce the enhancement effect in high-noise scenarios; (2) the computational complexity of the CNN-LSTM-Attention model is high, which is not applicable to real-time monitoring scenarios with limited resources ; (3) the interpretability of model-generated data is weak, and the adaptability in special failure modes needs to be improved. In the future, the computational efficiency of the model can be considered to be enhanced by dimensionality reduction techniques and parallel processing. In terms of data preprocessing, methods such as variational self-encoder can be tried to enhance the robustness of the model and strengthen the ability to handle outliers and missing data.

Acknowledgment

The authors are grateful to the anonymous reviewers, and the editor for their critical and constructive review of the manuscript. This study was co-supported by the National Natural Science Foundation of China (51507074), Postgraduate Research &Practice Innovation Program of Jiangsu Province (KYCX25_4368).

References

1. Wang, Z., Li, W., Shang, S., Wang, Z., Han, C. (2020). Performance degradation comparisons and failure mechanism of silver metal oxide contact materials in relays application by simulation. Eksploatacja i Niezawodność – Maintenance and Reliability, 22(1), 86-93. https://doi.org/10.17531/ein.2020.1.10.
2. Wang X., Wang BX., Wu W., Hong Y. Reliability analysis for accelerated degradation data based on the Wiener process with random effects. Quality & Reliability Eng 2020;36:1969–81. https://doi.org/10.1002/qre.2668.
3. Liu, Y., Wang, Y., Fan, Z., Hou, Z., Zhang, S., Chen, X. (2020). Lifetime prediction method for MEMS gyroscope based on accelerated degradation test and acceleration factor model. Eksploatacja i Niezawodność – Maintenance and Reliability, 22(2), 221-231. https://doi.org/10.17531/ein.2020.2.5.
4. Su, Z., Hua, Z., Hu, D., Zhao, M. (2025). Research on Equipment Reliability Modeling and Periodic Maintenance Strategies in Dynamic Environment. Eksploatacja i Niezawodność – Maintenance and Reliability, 27(1). https://doi.org/10.17531/ein/192163.
5. Wang, Z., Shang, S., Wang, J., Huang, Z., & Sai, F. (2017). Accelerated storage degradation testing and failure mechanisms of aerospace electromagnetic relay. Eksploatacja I Niezawodnosc-maintenance and Reliability, 19, 530-541. https://doi.org/ 10.17531/ein.2017.4.6.
6. Zhan, X., Liu, Z., Yan, H., Wu, Z., Guo, C., Jia, X. (2023). A novel method of health indicator construction and remaining useful life prediction based on deep learning. Eksploatacja i Niezawodność – Maintenance and Reliability, 25(4). https://doi.org/10.17531/ein/171374.
7. Wang, Z., Shangguan, W., Peng, C., Cai, B. (2024). Similarity Based Remaining Useful Life Prediction for Lithium-ion Battery under Small Sample Situation Based on Data Augmentation. Eksploatacja i Niezawodność – Maintenance and Reliability, 26(1). https://doi.org/10.17531/ein/175585.
8. Lu, Y., Shen, M., Wang, H., Wang, X., van Rechem C., Fu T., et al. Machine Learning for Synthetic Data Generation: A Review 2024.
9. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680). https://doi.org/10.1109/ICCCNT56998.2023.10306417.
10. Yoon, J., Jarrett, D., & Schaar, M.V. (2019). Time-series Generative Adversarial Networks. Neural Information Processing Systems.

11. Yang, L., He, P., & Chou, X. (2024). TimeGAN data augmentation-based fault classification method for complex processes. Journal of Zhejiang University (Engineering Science), 59(9), 1768–1780. https://doi.org/10.3785/j.issn.1008-973X.2024.09.002.
12. Ma, Z. , Sun, Y. , Ji, H. , Li, S. , Nie, S. , & Yin, F. (2024). A CNN-BiLSTM-Attention approach for EHA degradation prediction based on time-series generative adversarial network. Mechanical Systems & Signal Processing(Jun.), 215. https://doi.org/10.1016/j.ymssp.2024.111443.
13. Chen, Y., Wang, X., Wang, Y., Gao, Y., Qu, J., Dai, H., & Xu, C. (2024). Multi-combination fault data augmentation method of aero-engine gas path based on Extraction TimeGAN. Measurement, 225, 115778. https://doi.org/10.1016/j.measurement.2024.115778.
14. Z. G. Liu, T. Y. Ji, J. W. Chen, L. J. Zhang, L. L. Zhang and Q. H. Wu, "Conditional-TimeGAN for Realistic and High-Quality Appliance Trajectories Generation and Data Augmentation in Nonintrusive Load Monitoring," in IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1-15, 2024, Art no. 2514415. https://doi.org/10.1109/TIM.2024.3381263.
15. W. Deng *et al*., "Short-Term Wind Power Prediction Based on Wind Speed Interval Division and TimeGAN for Gale Weather," 2023 International Conference on Power Energy Systems and Applications (ICoPESA), Nanjing, China, 2023, pp. 352-357. https://doi.org/10.1109/ICoPESA56898.2023.10140676.
16. Z. Zhi., L. Liu, D.. Liu and C. Hu, "Fault Detection of the Harmonic Reducer Based on CNN-LSTM With a Novel Denoising Algorithm," in IEEE Sensors Journal, vol. 22, no. 3, pp. 2572-2581, 1 Feb.1, 2022. https://doi.org/10.1109/JSEN.2021.3137992.
17. Xu Y., Xiang Y., &Ma T.(2022). Short-term Power Load Forecasting Method Based on EMD-CNN-LSTM Hybrid Model. Journal of North China Electric Power University(02), 81-89. https://doi.org/10.3969/j.ISSN.1007-2691.2022.02.09.
18. Hou H., Wu W., Wei R., HEe H., Wang L., Li Z & Lin X.(2024). Attention Mechanism Based CNN-LSTM-XGBoost Electric Power Meteorological Hybrid Forecasting Model of Typhoon Rainstorm. Smart Power & Energy Security (10), 96-102. https://doi.org/10.20204/j.sp.2024.10013.
19. Chu Y., Chen Y & Mi Y.(2022). Lithium battery health state estimation based on attention mechanism CNN-LSTM. Chinese Journal of Power Sources (06), 634-637+651.https://doi.org/10.3969/j.issn.1002-087X.2022.06.013.
20. J. Liu, H. Gao, Q. Yan, B. Bai and X. Wang, "Short-term Load Forecasting Method Based on LSTM-Attention Mechanism," *2021* IEEE Sustainable Power and Energy Conference (*iSPEC)*, Nanjing, China, 2021, pp. 3837-3842 .https://doi.org/10.1109/iSPEC53008.2021.9735593.
21. Zhongshan, H., Ling, T., Dong, X., & Yaozhong, W. (2018). Prediction of oil temperature variations in a wind turbine gearbox based on PCA and an SPC-dynamic neural network hybrid. Journal of Tsinghua University, 58, 539-546. https://doi.org/10.16511/j.cnki.qhdxxb.2018.26.026.