Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P3423-cd

Remaining useful life prediction for train bearing based on an BiLSTM-KAN

Yiwei Zheng

School of Automation and Intelligence, Beijing Jiaotong University, China. E-mail: 23125152@bjtu.edu.cn

Fei Yan

Urban Rail Transit Beijing Laboratory, National Engineering Research Centre of Rail Transportation Operation and Control Systems, School of Automation and Intelligence, Beijing Jiaotong University, China. E-mail: fyan@bjtu.edu.cn

As one of the key components of the train bogie, accurate bearing remaining useful life (RUL) prediction and timely maintenance play a vital role in the safe and reliable operation of the train. The environment of rail transit trains is complex, and the vibration signal of train bearings shows the characteristics of non-linearity and non-smoothness. Meanwhile, the safety requirements of rail transportation system are comparatively demanding, and the time series RUL prediction of bearings should consider the long-term and multi-data problems. For the complex degradation process of rail transit train bearings, a hybrid bidirectional long and short-term memory (BiLSTM) networks and Kolmogorov-Arnold Networks (KAN) RUL prediction method is proposed. Based on the BiLSTM network, KAN is used to replace the fully connected layer, which improves the parameter utilization and enhances the ability to obtain the nonlinear pattern information in the hidden state of BiLSTM. Compared with the traditional time-series prediction method, the method has better prediction accuracy, stronger interpretability, and is more suitable for the prediction of train bogie bearing RUL in high safety requirement scenarios.

Keywords: Train bearing, RUL, BiLSTM, KAN.

1. Introduction

China's rail network has developed rapidly in recent years. China's high-speed rail network has exceeded 45,000 kilometres. The reliability and safety of railway trains have received great attention from all walks of life. At present, regular preventive maintenance is a common maintenance method in train operation management, but there are some problems (Dai et al., 2023). Over-maintenance increases costs, and under-maintenance can lead to equipment failure, downtime or even endanger personal safety. Therefore, Prognostics and Health Management (PHM) is widely used and RUL is an important basis for assessing the health condition (Wang et al., 2021). There is no doubt that the performance of bogie bearings is directly related to these priority objectives. Ensuring that bogie bearings are in good condition to avoid unplanned downtime or catastrophic failure is critical and can significantly reduce costs and improve operator profitability by extending bearing RUL. Predictive or condition-based maintenance strategies are widely used to monitor bearing health in real time, extend actual bearing RUL, prevent catastrophic failures and save operating costs. On-board condition monitoring and bogie bearing health diagnosis or prediction are essential in train operations.

The physical model-based method is predicated on the measurement of the object in order to establish an accurate mechanism model. This is then compared with the actual output of the degradation mechanism model. Mathematical methods are then employed to analyse and process the residuals, thus achieving the desired level of degradation (Luo et al., 2008). The degradation of train bogie bearings is affected by a dynamic and complex operating environment, making it difficult to develop highly accurate predictive models. The artificial neural network approach, a data-driven method, builds a historical degradation model using equipment data and monitors real-time degradation levels through this model. It is easy to implement, requires no prior knowledge or physical models, and relies solely on processing and analyzing operating data. [4]. The use of recurrent neural networks (RNN) in analysing time series data has proven to be a valuable tool for extracting temporal and dynamic features from sequences, leading to its widespread application in RUL. (Guo et al., 2017) proposed a method that combined six correlation-similarity features and eight classical time-frequency features, selecting the most sensitive features using monotonicity and correlation indexes, and constructing the RNN-HI through the use of a RNN. However, the degradation of train bogie bearings is a long-term process due to the high safety guarantee, and long-term data need to be processed in the actual process. The LSTM variant of the RNN is a suitable solution to this long-term dependency problem. (Liu et al., 2021) analysed multi-stage bearing degradation through a statistical process, which helps LSTM to make predictions, but the generalisation ability is weak. (Cheng et al., 2021) proposed a new method for rolling bearing RUL prediction based on convolutional neural network (CNN) and BiLSTM models, which was trained by constructing nonlinear degradation indices (DI) to achieve an accurate prediction of future degradation indicators and RUL of rolling bearings. In train bogie bearing RUL prediction, data-driven methods provide high accuracy and flexibility but face significant drawbacks. They heavily depend on the quality and quantity of training data, leading to performance degradation with incomplete or noisy data. Additionally, their complexity demands substantial computational resources, resulting in high time and economic costs.In light of the aforementioned limitations, a novel approach for RUL prediction of train bogie bearings is proposed by combining the BiLSTM algorithm with the KAN (Liu et al., 2024). The LSTM algorithm has been demonstrated to excel at capturing long-term dependencies in sequential data.The KAN component has been shown to be used to efficiently map complex nonlinear offsets.Experiments by (Liu et al., 2024) have demonstrated that KAN is more efficient than multilayer perceptron (MLP) in terms of parameter utilisation, and can achieve better fitting with fewer nodes and less training time. The integration of these two approaches has been shown to enhance the extraction of features and the recognition of time-series information, leading to a substantial enhancement in the accuracy of predicting the RUL of train bogie bearings.The primary contributions of this paper are outlined as follows:

(i) The process of monitoring the degradation of train bogie bearings is accomplished through the utilisation of a BiLSTM network in conjunction with a KAN. The KAN is employed to extract key trends from the sequence for fitting.
(ii) Multiple degradation features were constructed to establish a framework for train bogie bearing RUL prediction based on BiLSTM-KAN.
(iii) A comprehensive and detailed validation of the BiLSTM-KAN method was carried out, mainly on the XJTU-SY dataset, to confirm the efficiency of the method, and a comparative evaluation of multiple methods using R^2, RMSE, and MAE. The validity of the model was further validated.

The rest of the paper is organised as follows: section II describes the proposed BiLSTM-KAN prediction framework. The LSTM theory and KAN theory are also explained. In section III, the effectiveness of the prediction framework is verified by an experimental study on the XJTU-SY dataset. It includes model evaluation indexes, prediction results, multi-model comparison, and result analysis. Finally, in section IV, conclusions are given and possibilities for future research are discussed.

2. Model Construction

2.1.*Feature extraction*

To improve bearing RUL prediction, it is essential to extract enhanced characteristic indicators from monitoring signals that accurately reflect train bearing degradation. From bearing vibration acceleration signals, time-domain and frequency-domain features can be extracted. Time-domain analysis describes the temporal variation of vibration signals, with features like crest factor and skewness providing valuable degradation time information for long-term dependency prediction. Frequency-domain analysis detects anomalous shocks and describes bearing conditions, using parameters such as kurtosis, entropy value, and energy ratio to identify variations in the frequency spectrum.

2.2. *Network structure*

2.2.1. *Long short-term memory network*

LSTM is a form of RNN that can effectively capture long term relationships in time series data. LSTM network is based on RNN and introduces three threshold structures in the cell named as input gate, forget gate and output gate to selectively discard, determine, update and output information (Yan et al., 2022).

Discarding cell state information is the initial action of LSTM. The forget gate receives the hidden state from the previous time step and the input from the current time step. It then outputs a value between 0 and 1 using the Sigmoid activation function. A value close to 1 means that more information will be retained and a value close to 0 means that more information will be forgotten, as shown in Eq.(1):

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

In this case, the output of the forget gate is denoted by f_t. The input of the current time step is x_t. The weight matrix and bias vector to be learnt are W_f and b_f, respectively. the hidden state of the previous LSTM unit is h_{t-1}.

The input gate determines what new information is added to the cell state. It is also activated using the Sigmoid function and combined with the new candidate memory cell state as shown in Eq.(2):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

The output of the input gate is denoted by i_t. The weight matrix and bias vector to be learnt are W_i and b_i respectively.

The tanh function is then used to construct a new vector $\tilde{C}_t$ and add it to the unit state.

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

The weight matrix and bias vector to be learnt are W_c and b_c.

Updating the state of a memory cell requires a combination of forget and input gates. The cell state C_t is updated by the following Eq.(4):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

Ultimately, the output o_t and the current hidden state h_t are regulated by the output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$
$$h_t = o_t \cdot tanh(C_t) \quad (6)$$

For time series problems, the accuracy of prediction is not only related to the information in the previous moment, but also to the information after the current moment. LSTM is a unidirectional transmission in which only the information in the previous moment is taken into account in the state transmission process. Bi-LSTM is a bidirectional transmission, which includes both a forward layer and a backward layer, and links the states of the forward layer and the states of the backward layer to the same output layer, so that the output can take into account both past and future state information, thus improving the prediction accuracy.

At each time step t, the hidden states h_t^f and h_t^b of the forward LSTM and the reverse LSTM are spliced together to form the final bidirectional hidden state, and the output of the network can be represented as follows:

$$h_{(t)} = \left[h_t^f, h_t^b\right] \quad (7)$$

The structure of BiLSTM is shown in Fig.1.

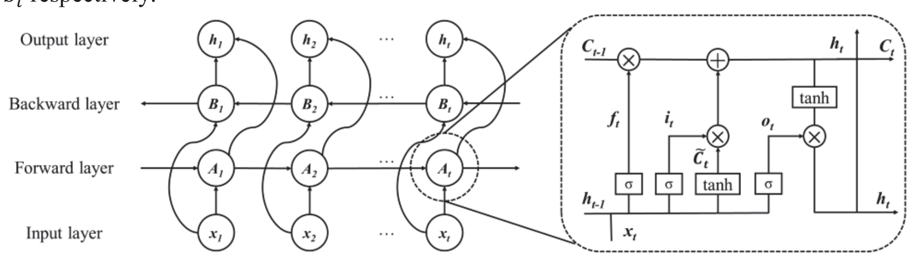

Fig. 1. The structure of BiLSTM

2.2.2. *Kolmogorov-Arnold networks*

Kolmogorov-Arnold Networks (KAN) are a novel neural architecture inspired by the Kolmogorov-Arnold Representation Theorem. Unlike MLPs (Tolstikhin et al., 2021) that apply fixed node activations after summation, KAN places learnable activation functions (e.g., spline-based) on edges between nodes, implementing a "nonlinearity-first, summation-later" paradigm. This design enables KANs to achieve greater flexibility and parameter efficiency while dynamically adapting activation patterns during training. By jointly optimizing connection weights and edge-specific mappings, KANs excel at extracting complex nonlinear features, particularly in applications like bearing vibration signal analysis where intricate data patterns demand adaptive function learning.

The Kolmogorov-Arnold representation theorem, which is the basis of KAN theory, states that any continuous function can be written as a combination of a set of unitary functions.

For any continuous function f, there exists a set of unitary functions Φ_q and $\phi_{q,p}$ satisfying Eq.(8)

$$f(x_1, \dots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right) \tag{8}$$

The unitary functions ϕp and ψ are represented and learnt by b-splines. A Spline Function is a smoothing function for approximating or interpolating data, which consists of segmented polynomials spliced together and these polynomials have a certain degree of continuity at the junction.B-Splines, on the other hand, represent a spline using a set of basis functions, each of which is non-zero in only a few subintervals. This method can be used to improve local accuracy by adjusting the density of control points.

In MLPs, once a layer (consisting of linear transformations and nonlinearities) is defined, more layers can be stacked to make the network deeper. A KAN layer with input and output dimensions can be defined as a one-dimensional function matrix:

$$\Phi = \{\phi_{q,p}\},\ p = 1,2, \dots, n_{in},\ q = 1,2, \dots, n_{out} \tag{9}$$

where the function $\phi_{q,p}$ has trainable parameters.

The shape of a KAN is represented by an array of integers $[n_0, n_1, \dots n_L]$, where n_i is the number of nodes in the i^{th} layer of the computational graph. We denote the i^{th} neuron in the l^{th} layer by (l, i), and the activation value of the (l, i) neuron by $x_{l,i}$. Between layer l and layer $l + 1$, there are $n_l n_{l+1}$ activation functions: the activation function that connects (l, i) and $(l + 1, j)$ is denoted by Eq.(10)

$$\phi_{l,j,i},\quad l = 0, \dots, L-1,\ i = 1, \dots, n_l,\ j = 1, \dots, n_{l+1} \tag{10}$$

The preactivation of $\phi_{l,j,i}$ is simply $x_{l,i}$, the postactivation of $\phi_{l,j,i}$ is denoted by $\tilde{x}_{l,j,i} \equiv \phi_{l,j,i}(x_{l,i})$, and the activation value of the $(l + 1, j)$ neuron is simply the sum of all incoming postactivations

$$x_{l+1,j} = \sum_{i=1}^{n_l} \tilde{x}_{l,j,i} = \sum_{i=1}^{n_l} \phi_{l,j,i}(x_{l,i}), j = 1, \dots, n_{l+1} \tag{11}$$

In matrix form, this can be written as

$$\mathbf{X}_{l+1} = \begin{pmatrix} \phi_{l,1,1} & \phi_{l,j,i} & \cdots & \phi_{l,j,i} \\ \phi_{l,j,i} & \phi_{l,j,i} & \cdots & \phi_{l,j,i} \\ \vdots & \vdots & & \vdots \\ \phi_{l,j,i} & \phi_{l,j,i} & \cdots & \phi_{l,j,i} \end{pmatrix} \mathbf{X}_l \tag{12}$$

$$\mathbf{X}_{l+1} = \boldsymbol{\Phi}_l \mathbf{X}_l \tag{13}$$

Where $\boldsymbol{\Phi}_l$ is the function matrix corresponding to the l^{th} layer (B-spline function matrix) and x is the input matrix.

A general KAN network is composed of L layers: given an input vector $\mathbf{x_0} \in \boldsymbol{R}^{n_0}$, the output of the KAN is

$$\text{KAN}(\text{x}) = (\boldsymbol{\Phi}_{L-1} \circ \boldsymbol{\Phi}_{L-2} \circ \cdots \circ \boldsymbol{\Phi}_1 \circ \boldsymbol{\Phi}_0)\text{x} \tag{14}$$

The simplest KAN can then be written as: $f(x) = \boldsymbol{\Phi}_{out} \circ \boldsymbol{\Phi}_{in} \circ x$, as shown in Fig 2.

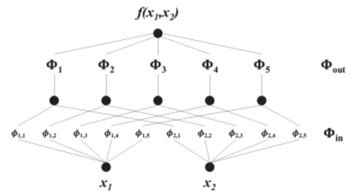

Fig. 2. The structure of KAN

In this way, each KAN connection is equivalent to a ‘mini-network’, enabling greater expressive power.

2.3. *Explainability train bearing RUL prediciton*

Combining the advantages of BiLSTM in processing time series tasks with the expressive ability and parameter utilisation efficiency of KAN, the BiLSTM-KAN integrated neural network model is proposed, as shown in Fig. 3.

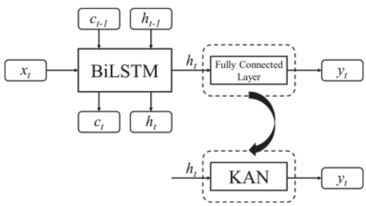

Fig. 3. BiLSTM-KAN

The present study proposes the utilisation of KAN as a replacement for the fully connected layer through which BiLSTM originally passes. The KAN network employs B-Splines as the activation function, facilitating the learning of the correct univariate function. Through visualisation, KAN can be interpreted, thereby enabling the observation of the thinking process and the internal operation mechanism of the model. It can reveal the composition structure and variable dependency of the synthetic data set through symbolic formulas, successfully fit symbolic expressions close to the data generating formulas, and directly show the mathematical relationship between input variables and outputs, realizing interpretability based on symbolic formulas, which is difficult to achieve with traditional neural networks. The field of rail transportation, as a representative of high safety, has high requirements for the interpretability of the methods used, which can ensure the transparency and fairness of the work process and help to divide responsibility after an accident. The use of KAN network as the decision-making output of the model can effectively improve the interpretability of the model, and the visualization process of KAN can better help researchers to analyze the decision-making process and validate it to ensure the safe operation of trains.

3. Experiment Analysis

3.1. *Bearing vibration signal data*

Due to the high safety of rail transit system operation, it is difficult to collect the whole life cycle vibration signal data of train bearings at the actual engineering site. In this paper, the publicly available XJTU-SY bearing dataset is used for experiments (Wang et al., 2018). Testbed of rolling element bearings, shown in Fig. 4, consists of an AC motor, motor speed controller, hydraulic loading system and test bearings, etc. It can carry out accelerated life tests of various types of rolling bearings or plain bearings under different working conditions, and obtain the full-life cycle monitoring data of the test bearings. The test is designed for three types of working conditions, 1) load 12 kN, speed 2100 rpm; 2) load 11 kN, speed 2250 rpm; 3) load 10 kN, speed 2400 rpm. 25,600 points are generated by sampling once per minute. Each sample was taken for 1.28 s, yielding 32,768 data points.

The rotational speeds under the three working conditions are comparable to those under the high-speed operating condition of rail transit trains, which can simulate the full life cycle state of bearings under the high load operating condition of trains. And in the actual process, there is consistency in the deterioration trend failure characteristics of the bearings, so the experimental results using this data set are representative.

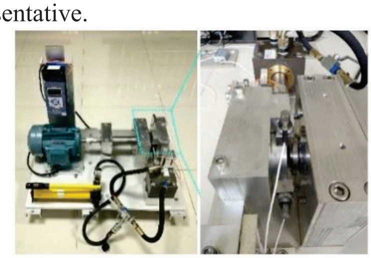

Fig. 4. Testbed of rolling element bearings

3.2. *Data processing*

For degraded faulty bearings, 13 time and frequency domain features are extracted from the life cycle data as shown in Section 2.1. Subsequently, z-score normalisation is used as the normalised model (), which is divided into training, validation and test sets after processing through a sliding window. The actual bearing RUL is considered as a linear degradation and is used as a label for the training and test sets. The initial health index of the bearing is defined as 1, and the health index when running to failure is defined as 0.

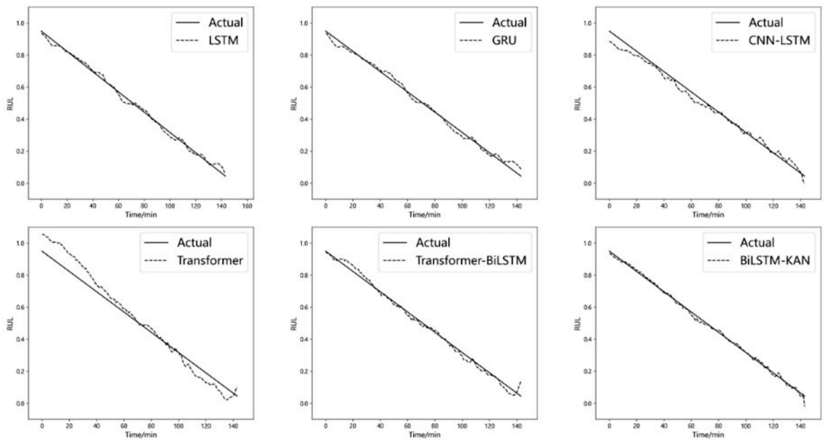

Fig. 5. RUL prediction results of Bearing 1_3 by several models

3.3. *Results of RUL prediction*

The predicted results for bearings 1_3 are shown in Figure 5.

From the prediction curves, LSTM and GRU (Ravanelli et al., 2018) were able to basically fit the actual curves.CNN-LSTM fitted poorly at the beginning and achieved better results after accumulating long time data. Transformer (Liu et al., 2021) performed poorly on the actual curves, with high predicted lifetimes at the initial stage and low predicted lifetimes at the later stage, which made the bearings fail early and failed to be fully functional. Transformer-BiLSTM was able to achieve a better fit to the actual curve, however, it did not converge at the final time period, which did not meet the practical requirements.The predicted curve of the BiLSTM-KAN model almost completely overlapped with the curve of the actual RUL value and converged before the end of the life, which was somewhat safe. This indicates that the model is highly accurate in capturing the trend of RUL changes and is more in line with the actual RUL.

In order to further verify the feasibility and generalisation ability of BiLSTM-KAN to predict RUL, experiments were conducted on rolling bearings under different operating conditions. The test bearings have degradation faults such as bearing 1_3, bearing 1_4, bearing 2_5 and bearing 3_4. The distributions are shown in Table 1.

Table 1. Bearing feature

Bearing	Fault	Working Condition
Bearing1_3	Outer	2100rpm 12kN
Bearing1_4	Cage	
Bearing2_5	Outer	2250rpm 11kN
Bearing3_4	Inner	2400rpm 10kN

The RUL prediction results of BiLSTM-KAN are shown in Fig. 6.

The predicted curves from these bearing datasets show a good fit to the actual bearing life curves in the vast majority of cases. For Bearing2_5, Bearing3_4 data volume larger longer time period bearings, there is a certain oscillation in the late prediction, when the bearing is in the serious failure stage, there is a certain impact on the life prediction of the bearing, but the final complete failure threshold is correctly converged. In the stable wear stage of the bearing, the predicted value is very close to the true value, and the surface of the proposed model has good performance and can cope with different working conditions and express the severe failure of the bearing.

In order to evaluate the performance of the proposed model and other models, regression metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination R-Square (R2) were used to evaluate the prediction performance (Guo et al., 2021). And it was compared with other optimisation models in the experiment. The results of the metrics for the above four bearings are shown in Table 2.

As can be seen from Table 2, the performance metrics of BiLSTM-KAN are improved compared to LSTM, GRU and Transformer-BiLSTM.

The average RMSE is reduced by more than 18.6% and the MAE is reduced by more than 26.72%, indicating that the proposed BiLSTM-KAN has a low prediction error.The R2 is improved by at least 1.62% on average, which indicates that the regression performance of BiLSTM-KAN is better. The analysis results show that KAN replaces the fully connected layer to improve the prediction performance of BiLSTM.

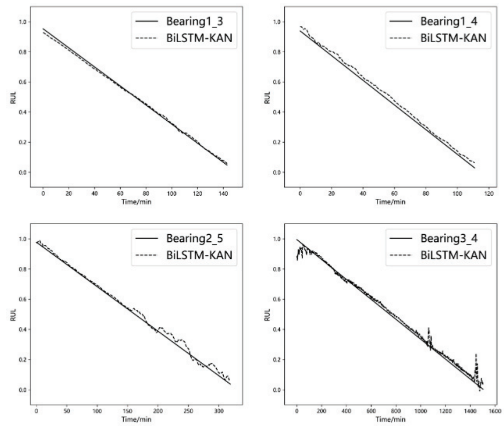

Fig. 6. RUL prediction results of BiLSTM-KAN for bearings under different working conditions

Table 2. RUL prediction performance indexes

Bearing	Index	LSTM	GRU	Transformer-BiLSTM	BiLSTM-KAN
Bearing1_3	RMSE	0.02088	0.02010	0.04024	**0.01133**
	MAE	0.01759	0.01741	0.03181	**0.00982**
	R^2	0.99370	0.99416	0.97660	**0.99814**
Bearing1_4	RMSE	0.02558	0.03330	0.04588	**0.01231**
	MAE	0.01991	0.02408	0.03400	**0.01014**
	R^2	0.99069	0.98421	0.97003	**0.99784**
Bearing2_5	RMSE	0.02450	**0.02013**	0.06078	0.02248
	MAE	0.02024	0.01396	0.05198	**0.01208**
	R^2	0.99191	0.99454	0.95025	**0.99620**
Bearing3_4	RMSE	0.03511	**0.02622**	0.05193	0.02869
	MAE	0.02351	0.01944	0.03146	**0.01929**
	R^2	0.98499	0.99163	0.96716	**0.99498**

4. Discussion

The BiLSTM-KAN model demonstrates superior RUL prediction accuracy for train bearings, validated by lower RMSE/MAE and higher R^2 values (Table 2). BiLSTM's bidirectional architecture captures both historical and future degradation trends, enabling early-stage accuracy (Figure 5). KAN's B-spline-based activation functions enhance nonlinear fitting while reducing parameter redundancy, critical for embedded rail systems. For example, BiLSTM-KAN achieves a 45.7% lower RMSE than LSTM for Bearing1_3, highlighting its efficiency in handling noisy vibration signals. Unlike CNN-LSTM, which requires extensive data accumulation, BiLSTM-KAN provides stable predictions from initial degradation phases. Transformer-BiLSTM struggles with late-stage convergence due to sparse attention patterns, whereas KAN's localized B-spline adjustments ensure robustness in severe failure scenarios. KAN's interpretability—visualizing feature weights via spline functions—addresses the "black-box" distrust in safety-critical applications, aiding fault attribution.

While the BiLSTM-KAN model shows promise, its reliance on the XJTU-SY dataset limits generalizability due to a lack of real-world variability. KAN's B-spline optimization also increases training time by 15–20%, challenging real-time deployment. Late-stage prediction oscillations (e.g., Bearing2_5/3_4) suggest the need for noise suppression techniques like wavelet filters. Future work should focus on real-world validation, physicsbased model integration, and KAN optimization for edge deployment.

5. Conclusions

The article proposes a BiLSTM-KAN based network model for train bogie bearing RUL prediction, which is of positive significance for the safe operation of trains. The method uses BiLSTM to extract feature information and capture long-term dependencies in bearing time-series data, and connects the KAN layer to further access the information in the hidden states through a learnable nonlinear activation function, thus improving the model's ability to fit the data.

In addition, our team is in the process of refining the dataset using our own bearing fault test platform for building simulated train bogies(Fig. 7). In the future, we intend to use our own dataset, which is more related to trains, to validate and test the model proposed in this paper, to improve the performance of the model and to better ensure the safe operation of trains.

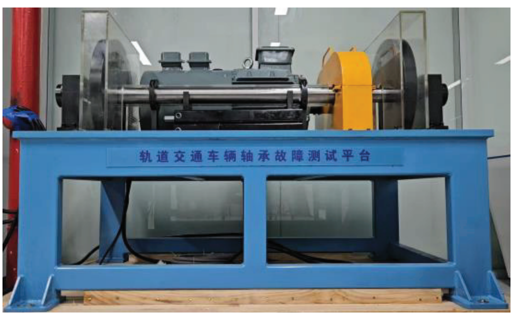

Fig. 7. Train bearing fault test platform

Acknowledgement

This work was supported by National Natural Science Foundation of China Regional Innovation Joint Fund Key Project (U22A2053), Guangxi Key RD Project (Gui scientific AB23075209)

References

Cheng, Yiwei, et al. (2021). "A Convolutional Neural Network Based Degradation Indicator Construction and Health Prognosis Using Bidirectional Long Short-Term Memory Network for Rolling Bearings.",Advanced Engineering Informatics 48

Dai, Jisheng, et al. (2023). "Multi-component System Maintenance Optimization of a Rail Transit Train Based on Opportunistic Correlations", TRANSPORTATION SAFETY AND ENVIRONMENT 5.4

Fei, Nanyi, et al. (2021). "Z-Score Normalization, Hubness, and Few-Shot Learning.",IEEE International Conference on Computer Vision

Guo, Jianwen, et al. (2021). "Mechanical fault time series prediction by using EFMSAE-LSTM neural network",Measurement: Journal of the International Measurement Confederation 173

Guo, Liang, et al. (2017). "A Recurrent Neural Network Based Health Indicator for Remaining Useful Life Prediction of Bearings",Neurocomputing 240

Han, Wang, et al. (2021). "Remaining Useful Life Prediction and Optimal Maintenance Time Determination for a Single Unit Using Isotonic Regression and Gamma Process Model", Reliability Engineering & System Safety 210: 107504-107504.

Liu, Junqiang, et al. (2021). "Fault Prediction of Bearings Based on LSTM and Statistical Process Analysis.",Reliability Engineering & System Safety 214

Liu, Ze, et al. (2021). "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows",IEEE International Conference on Computer Vision

Liu, Ziming, et al. (2024). "KAN: Kolmogorov-Arnold Networks",CoRR abs/2404.19756

Luo, Jianhui et al. (2008). "Model-Based Prognostic Techniques Applied to a Suspension System", AUTOTESTCON 38.5

Ravanelli, Mirco, et al. (2018). "Light Gated Recurrent Units for Speech Recognition",IEEE Transactions on Emerging Topics in Computational Intelligence 2.2

Saidi, Lotfi, et al. (2017). "Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR", Applied Acoustics 120.

Tolstikhin, Llya, et al. (2021). "MLP-Mixer: an All-Mlp Architecture for Vision",Conference on Neural Information Processing Systems 34

Wang, Biao, et al. (2020). "A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings.",IEEE Transactions on Reliability 69.1

Yan, Guangxi, et al. (2023). "A Hybrid Ensemble Deep Reinforcement Learning Model for Locomotive Axle Temperature Using the Deterministic and Probabilistic Strategy",Transportation Safety and Environment 5.3

Zhu, Jun, et al. (2019). "Estimation of Bearing Remaining Useful Life Based on Multiscale Convolutional Neural Network.",IEEE Transactions on Industrial Electronics 66.4