Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 6
ISET 2024

Intelligent prediction for remaining life of aero-engine

--A Study on CNN-BiLSTM Model Based on Scores Loss Function

Yufei Li[#, 1, *], Yujie Xi[#, 2], Xin Zhou[#, 3]

[1]Department of Applied Statistics, Northwest University, Xi’an, China, 710127

[2]Department of Accounting, Northwest University, Xi’an, China, 710127

[3]Department of Economic Statistics, Northwest University, Xi’an, China, 710127

*Corresponding author: 2021114014@stumail.nwu.edu.cn

#These authors contributed equally.

Abstract. With the advent of the era of big data and artificial intelligence, predicting and managing the remaining service life of aero-engines has become a key challenge, which has a direct impact on the enhancement of autonomous control capabilities. In this paper, a new deep learning model combining a one-dimensional convolutional neural network (1D-CNN) and a bidirectional long and short-term memory network (BiLSTM) is proposed to efficiently handle dynamic and uncertain engine health state data. The concurrent model takes full advantage of 1D-CNN in extracting local features and BiLSTM in capturing time-dependence, avoiding the loss in the information transfer process in conventional models. Empirical studies on NASA's C-MAPSS dataset show that the model significantly improves the accuracy and robustness in predicting the remaining service life, reducing the RMSE by up to 5.05% and the Scores by up to 54.89% compared to the conventional model. Especially in the short-term prediction task, the model shows higher stability and accuracy, which provides strong support for advance warning in the field of aviation safety.

Keywords: Aero-engine; Scores loss function; CNN; BiLSTM; Extreme random tree.

1. Introduction

As the wave of the era of big data and artificial intelligence rolls on, how to implement the strategy of aerospace power under new opportunities and challenges has become a hot issue to be solved in the modernisation of China. Based on this, this paper analyses the state of the aero-engine from a statistical point of view using massive data, and establishes a parallel neural network prediction model based on one-dimensional CNN and BiLSTM, which plays a key role in the data processing, health assessment, prediction and decision-making of the aero-engine, and predicts the remaining service life of the aero-engine with a high degree of accuracy, so as to increase the maintenance efficiency and benefits for the The core technology innovation and system safety operation in the aerospace field provides strong support.

In the rapid development of big data and artificial intelligence technology, the purpose of this paper is to combine deep learning technology to propose a new method of predicting the remaining service life of aero-engine, so as to make new progress in theory and practice.

In order to achieve the above purpose, this paper is based on the strategic policy of building a strong science and technology country and a strong aerospace country, combined with the data-driven approach, firstly, by finding and reading relevant literature, we clarify the theoretical basis of the remaining life prediction technology (RUL) and the advantages and disadvantages of the existing research methods, and put forward the problems; and then, we construct the RUL indexes of aero-engine prediction: by choosing the NASA's (United States Aviation and Aeronautics Administration) C-MAPSS dataset as the experimental data source, and in terms of data preprocessing and failure threshold definition, min-max data normalisation and sliding window technology are used to normalise and denoise the data, so as to use the extreme random tree method for feature extraction, and segmented linear RUL for the definition of failure thresholds; finally, this paper innovatively puts forward a method based on a one-dimensional convolutional neural network (CNN) with bipartite RUL. Neural Network (CNN) in parallel with Bidirectional Long and Short-Term Memory Network (BiLSTM), which ensures that the model can efficiently and accurately capture the spatial and time-series features of the aero-engine operating state, and apply the established model to predict the remaining life, and the results show that the model performs excellently on the dataset, and the result tends to be more accurate, especially in the prediction of the short remaining life case.

In addition, this paper also compares the prediction results of the proposed method with those of other related SOTA methods, and summarises the theoretical advantages and practical significance of the model, and proposes some future directions for in-depth research.

2. Review of Related Literature

From the existing papers, it can be seen that two main approaches have been developed for aero-engine RUL prediction: statistical model-based approach and data-driven approach.

2.1. Statistical model-based approach

The statistical model-based approach is dedicated to predicting the life of aircraft engines by building stochastic models.Wu et.al[1] established a nonlinear Wiener process model for predicting the RUL value of an aero-engine, and fused the multi-sensor monitoring data by using Kalman filtering algorithm.Li et.al[2] used a multivariate measurement function to represent the behaviour of the multi-sensor data, and then established a state-space model to accurately predict the RUL of an aero-engine.Cui et.al [3], on the other hand, proposed a comprehensive RUL prediction method based on time-varying particle filtering (TVPF), which adaptively selects the optimal state model through a sliding window and combines global/local information fusion techniques to estimate the RUL value. These methods not only predict the RUL values, but also provide the probability density function of the RUL values, thus providing more comprehensive information for maintenance and life assessment of aero-engines.

There are drawbacks in this approach, for example, the complexity of aircraft engine structure, the nonlinearity of the operating state, and the uncertainty of the degradation process, which add a lot of difficulties to the establishment of stochastic models. Therefore, it is difficult to build an accurate model to predict RUL effectively with the statistical model-based approach.

2.2. Data-driven approach

The data-driven approach mainly consists of constructing a prediction model to estimate the remaining service life of an engine by collecting and analysing a large amount of real-time monitored sensor data and applying technical means such as statistics and machine learning. Unlike traditional failure mechanism-based prediction methods, data-driven methods do not require professional knowledge and have better generalisation capabilities.

Many machine learning methods have been applied to data-driven prediction methods, such as Support Vector Regression (SVR), Hidden Markov Models (HMM), Support Vector Machines (SVM) and Extreme Learning Machines (ELM). For example, Li et.al [4] investigated an integrated deep multiscale feature fusion network for predicting the RUL of an aero-engine by manually selecting 14 sensors for feature extraction and fusion.Liu et.al [5] introduced clustering and a long short-term memory network (LSTM) for multi-stage prediction of the RUL values, and used seven sensor data as the key parameters.Guo et .al [6] combined empirical modal decomposition (EMD) with a correlation vector machine (RVM) to predict the RUL of an air turbine starter in an engine.

However, traditional machine learning methods require proper feature engineering including feature extraction and dimensionality reduction. Without the help of relevant expert knowledge, inappropriate features may lead to undesirable performance. Fortunately, deep learning algorithms are able to automatically extract high-level feature representations without the need for feature engineering, which has inspired more and more scholars to utilise deep learning methods for remaining life prediction.

Deep learning-based methods have attracted considerable attention in the field of residual life prediction, including Long Short-Term Memory Networks [7] (LSTMs) and Convolutional Neural Networks (CNNs).Xia et.al [8] conducted a study on RUL prediction using LSTM networks, and Xiang et.al [9] proposed an Attention-Ordered Neuron (AON-LSTM) based LSTM network to predict gear RUL.Liu et.al[10] used a physical model to improve the accuracy of LSTM in RUL prediction.Zhang et.al[11] fine-tuned a bi-directional LSTM (BiLSTM) network for the prediction of RUL under different operating conditions of an aero-engine and used all the sensor monitoring data for the model training. In contrast to previous forward-only LSTM studies, the bidirectional network presents each sequence in two different LSTMs: one that processes the information forward, and the other that processes it in reverse. In this way, the bidirectional LSTM network smoothes the predictions and reduces the effect of noise.

These methods were evaluated on an aero-engine degradation dataset and the experimental results showed the superiority of the bidirectional LSTM structure compared to some machine learning algorithms. Instead of LSTM-based deep learning methods, some studies have built a CNN-based model for RUL estimation and obtained similar results.CNNs are capable of extracting relevant local features and have excellent representation learning capabilities.CNNs have become a major solution in the field of RUL estimation.

In summary, in aero-engine RUL prediction, both time-series features and spatial-series features need to be considered. CNNs can extract spatial-series features, while LSTMs can extract time-series features. However, in mainstream schemes CNN and LSTM are serially connected, CNN extracts spatial sequence features and then inputs them to LSTM to extract time sequence features, this method causes large information loss and redundant computation.

Therefore, we adopt the parallel neural network in this paper, i.e., two different neural network structures, namely, one-dimensional CNN and bi-directional LSTM network, are connected together in parallel to form a composite model that works jointly and complements each other. Compared with the series structure, its main advantage is that it can extract and fuse the spatio-temporal features of the data in parallel and efficiently, maintain the fidelity of the information, achieve feature complementarity, enhance the model adaptability and computational efficiency, and especially perform better in the short remaining life prediction.

3. Theoretical Basis of Remaining Useful Life Prediction Technology

PHM (Prognostics and Health Management) technology uses sensor technology, data analytics and predictive modelling to achieve failure prevention and maintenance optimisation by monitoring, diagnosing and predicting the health status of equipment or systems. Its basic framework includes the steps of data acquisition, pre-processing and feature extraction, fault diagnosis and health assessment, remaining life prediction and maintenance decision-making, and performance monitoring and feedback optimisation. Through these steps, the PHM system continuously monitors the state of the equipment to improve reliability, safety and availability, thereby reducing maintenance costs and extending equipment life.

PHM technology helps to optimise equipment maintenance schedules and resource allocation by predicting and evaluating the Remaining Equipment Life (RUL).RUL, as a key output of PHM technology, enables effective management and extension of equipment life by monitoring and predicting the state of health of the equipment.

Aero-engines similarly utilise historical data to build a remaining life prediction database, and pre-processed operational process data is used for RUL prediction. However, with advances in engine sensing technology, data generation has accelerated, posing challenges in data transmission and security. To address these issues, the aero-engine is used as an edge device to construct a health index (HI) using edge computing. The compressed and extracted HI values are transmitted to a data analytics centre for RUL prediction and development of operational maintenance strategies. These strategies automatically trigger corresponding maintenance measures based on the set HI and RUL baseline values, such as four health levels of healthy, sub-healthy, faulty and failed, corresponding to different maintenance assurance strategies (maintenance, inspection, minor and major repairs). This automated operational maintenance programme can improve the maintenance efficiency and safety of aero-engines.

Deep learning methods have been explored in the remaining life prediction of aero-engines with some results. Common deep learning methods include: Supervised Learning, which mainly uses labelled training data to train the model, with the goal of learning the mapping between inputs and outputs; Unsupervised Learning, which uses unlabelled training data to learn the intrinsic structure and patterns of the data, with the goal of discovering hidden relationships in the data; and Reinforcement Learning, which uses a combination of an intelligent body and an aero-engine to predict the remaining life of an aero-engine. Unsupervised Learning, which uses unlabelled training data to learn the intrinsic structure and patterns of the data, with the goal of discovering hidden relationships in the data; Reinforcement Learning, which learns the optimal action strategy through the interaction of the intelligence with the environment, with the goal of maximising the cumulative rewards in a given environment; and Semi-Supervised Learning, which is trained using a small number of labelled and a large number of unlabelled samples, with the goal of using a large number of unlabelled samples for training. Semi-Supervised Learning, which uses a small number of labelled samples and a large number of unlabelled samples for training, with the goal of improving model performance using unlabelled data; and Transfer Learning, which reduces the need for a large number of labelled samples and improves the efficiency and performance of learning by transferring the learnt knowledge to a new task.

Combined with the introduction of the above deep learning techniques, considering that the remaining life prediction process of the aero-engine has the label "remaining life RUL value", it is supervised learning, which is suitable for processing the training data containing the label.

Further, considering that aero-engine data are mostly time series data and contain multiple features, it is necessary to consider both spatial and temporal dimensions. At the spatial level, considering that the engine monitoring data is a one-dimensional signal, the CNN convolutional neural network's excellent local feature extraction capability can be used to extract the main features of the aero-engine RUL; at the same time, the bidirectional LSTM is used to capture the temporal dependence in the whole sequence, so that the future information can be used to make predictions.

4. Construction of RUL indicators for aero-engine prediction

The data in this paper is derived from NASA's C-MAPSS dataset for simulating the degradation process of turbofan engines.C-MAPSS is a simulation toolkit designed to simulate commercial turbofan engines, allowing the user to set the parameters of the mission, the closed-loop controllers, and the flight conditions according to the requirements. Due to the complexity of aero-engine construction and the confidentiality of operational data, NASA used Commercial Modular Aero-Propulsion System Simulation software to generate this dataset, which is designed to test the performance of different models.

NASA researchers used the C-MAPSS toolbox to simulate degradation trajectories over a large number of full life cycles for specific models of turbofan engines. The dataset was divided into four sub-datasets based on operating conditions and failure modes and published on NASA's official website for use by academics developing and evaluating life prediction algorithms. In addition, NASA has launched a RUL prediction challenge called "PHM08", which uses data generated by C-MAPSS to make predictions, and participants are required to submit their predictions for comparative evaluation. The structure of a turbofan engine is shown schematically in Fig. 1, which includes components such as a fan, low-pressure compressor, high-pressure compressor, low-pressure turbine, high-pressure turbine, and combustion chamber, and it is a typical twin-rotor, large-containment-ratio turbofan engine.

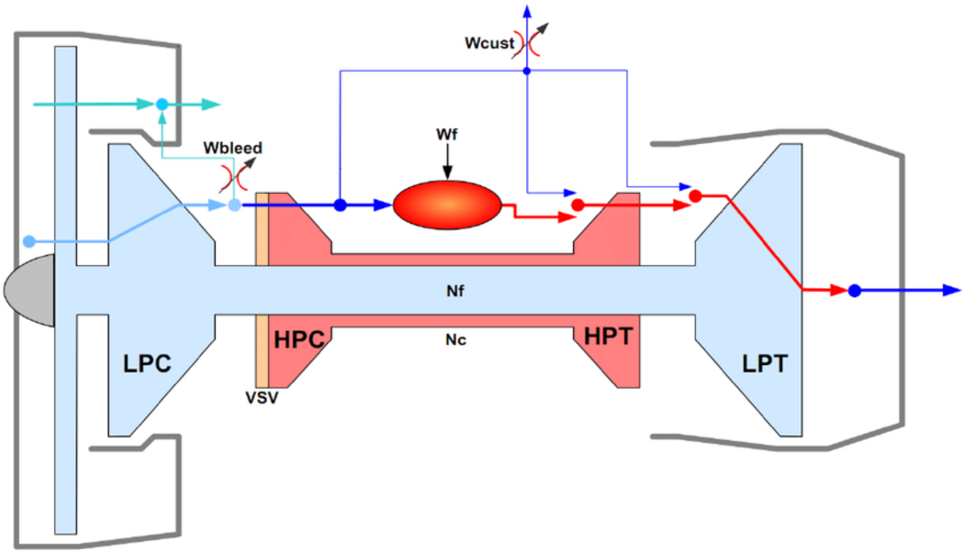

Figure 1 Simplified diagram of turbofan engine

Next, the data are preprocessed by min-max data normalisation followed by data denoising using sliding window averaging; then in terms of feature extraction, after initial screening using correlation heat map, the extreme random tree method is selected for further screening of the features so as to achieve the purpose of enhancing the model robustness and prediction accuracy; finally, the segmented linear RUL method is adopted to define the failure thresholds of the aero-engine. Finally, the segmented linear RUL method is adopted to define the failure threshold of the aero-engine, which is convenient for the subsequent RUL prediction.

5. Prediction method based on one-dimensional CNN-BiLSTM neural network

5.1. Determination of the asymmetric Scores-based loss function

In order to meet the safety requirements of RUL over prediction for aero-engines, this paper proposes an innovative asymmetric Scores loss function as the loss function for parallel model training. Compared with the traditional symmetric loss function, it can better cope with the stringent safety requirements in the aviation field and reduce the potential risks caused by the prediction lag.The definitions of Scores function and RMSE are shown in Eqs.(1) and Eq.(2), respectively.

$$S = \begin{cases} \sum_{i=1}^{n}(e^{-\frac{d_i}{13}} - 1), for\ d_i < 0 \\ \sum_{i=1}^{n}(e^{\frac{d_i}{10}} - 1), for\ d_i\ \geq 0 \end{cases} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {d_i}^2} \tag{2}$$

$d = \hat{t}_{RUL} - t_{RUL}$ (Forecast RUL - Actual RUL), n is the number of engines.

5.2. BiLSTM Principles of Bidirectional Long and Short Periodic Memory Neural Networks

Aero-engine remaining life prediction requires time series analysis of engine sensor monitoring data. These data usually contain complex time-series dependencies. To deal with such dependencies, the BiLSTM bidirectional long- and short-term memory neural network model is selected for this paper.

The LSTM network has a unique gating mechanism that captures and retains information over long time spans. This makes it highly capable of modelling long term trends and cyclical patterns in data sequences, which is crucial for understanding gradual changes during engine degradation. In contrast, traditional recurrent neural networks (RNNs) are prone to the gradient vanishing problem when dealing with long sequences, which makes it difficult for the model to learn long-distance dependencies.LSTM effectively solves the gradient vanishing problem and enables the model to maintain and update the far-away memories for long-distance dependencies that may be present in aero-engine data. For these two reasons, we chose the LSTM model.

BiLSTM, on the other hand, consists of two separate LSTM (Long Short-Term Memory Network) structures, one responsible for processing backwards from the start position of the input vector, called a forward LSTM, and the other responsible for processing forwards from the end position of the input vector, called a backward LSTM.The outputs of the two LSTM structures are connected together through a connectivity layer to form a comprehensive context vector.

5.3. Analysis and comparison of forecast results

5.3.1. Validation of the validity of the parallel model prediction: an example of engine 31 in FD001

Firstly, to verify the validity and superiority of the proposed parallel model prediction, we randomly selected a group of engine units from the FD001 test set for image illustration. As can be seen in Fig. 2, the predicted RUL values are close to the constant RULearly at the beginning.The RUL curve fitting results show that it decreases almost linearly until all the test samples are used up. Specifically, the early prediction error is greater than the late prediction error. This is due to the fact that more deterioration information becomes available as failures occur, resulting in higher prediction accuracy at the end of the process.

Figure 2 FD001 Dataset 31 Engine

5.3.2. Conclusion of remaining life prediction: an example of FD001 dataset

After completing the data processing and model building, the data are fed into the model for training. In this paper, we take the FD001 subset as an example and show the estimated and actual RUL results of the engine units in the test set, as shown in Figure 3.

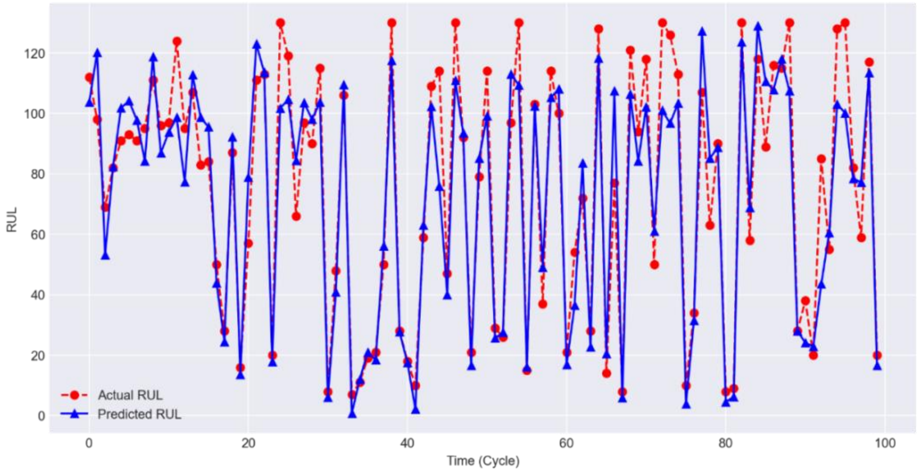

Figure 3 Engine RUL prediction for the test set

For a more intuitive analysis, the units are sorted by actual RUL value from smallest to largest, with the X-axis showing the actual RUL of the unit and the Y-axis indicating the predicted RUL, as shown in Figure 4.

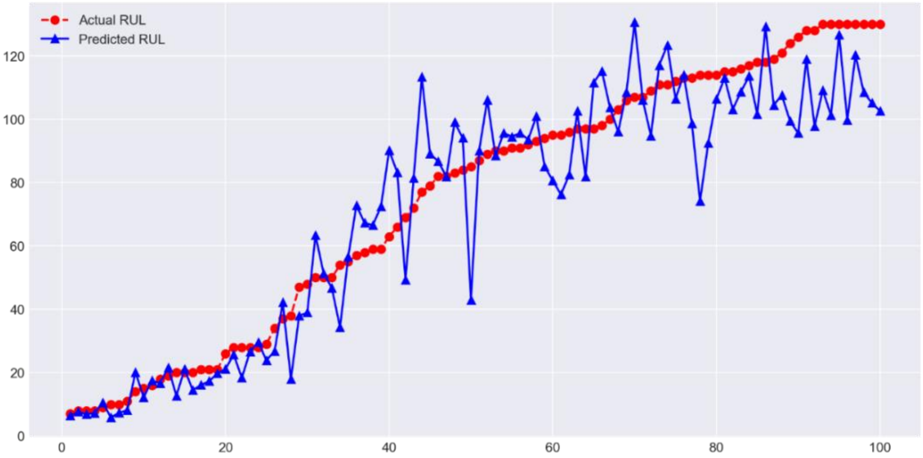

Figure 4 Predicted RUL of test engine with increasing RUL

It can be observed that the RUL prediction results tend to be more accurate when the actual RUL is lower. This is due to the fact that engine failure characteristics become more pronounced as the engine's service life comes to an end and the proposed methodology allows for more degradation information to be extracted from the sensor data. This is important in practice as it helps to accurately identify the risk of equipment failure and suspend the equipment in time to minimise losses.

5.3.3. Results of RUL prediction for all datasets

Meanwhile, this paper shows the estimated and actual RUL results for the test engine units in all four datasets. It can be concluded that the model we used has the best prediction results on the FD001 and FD003 datasets, followed by the FD002 dataset, and finally the FD004 dataset. Again, it can be seen that the RUL predictions tend to be more accurate when the actual RUL is lower.

5.3.4. Histogram of prediction error of the model

Figure 5, on the other hand, show the histograms of the prediction errors of our proposed model on the four subsets, where it can be observed that more errors are distributed to the left of 0, i.e., more underestimations than overestimations. This is caused by our choice to set the asymmetric Scores score function as the loss function.

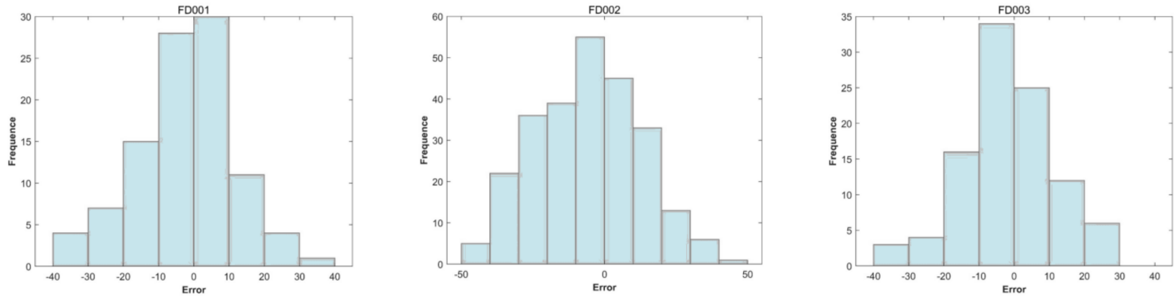

Figure 5 Histogram of prediction error for the dataset

5.4. Comparison with cutting-edge methods

The C-MAPSS dataset used in this paper is very popular in the field of remaining life prediction and many state-of-the-art (SOTA) frontier methods have been reported in recent years. Tables 1 and 2 show the RUL prediction performance of the proposed method and other frontier methods. In general, all methods show superior RMSE results for the FD001 and FD003 subsets compared to the results for the FD002 and FD004 subsets. These results stem from the multiple operating conditions and complex failure modes of FD002 and FD004, as well as differences in the size of the training dataset. As shown in Tables 1 and 2, our parallel model outperforms the SOTA model on the FD001, FD002, and FD003 subsets with RMSE reductions of 0.12, 1.08, and 0.30, which are equivalent to reductions of 0.95, 5.05, and 2.37 per cent, respectively.The Scores are reduced by 6.92, 1917.32, and 58.47, which are equivalent to reductions of 2.53%, 54.89%, and 20.59%. the reduction in Scores is significantly higher than the RMSE, thanks to our use of the score function as a loss function.

Table 1 RMSE results of the proposed method and related SOTA methods

methodologies	FD001	FD002	FD003	FD004
CNN	18.45	30.29	19.82	29.16
LSTM-FNN	16.14	24.49	16.18	28.17
DCNN	12.61	22.36	12.64	23.31
CatBoost	15.8	21.4	16.0	22.4
RNN	14.57	23.20	14.92	28.72
Proposed	12.49	20.32	12.34	26.54

Table 2 Scores results of the proposed method and related SOTA methods

methodologies	FD001	FD002	FD003	FD004
CNN	1287	13570	1596	7886
LSTM-FNN	338	4450	852	5550
DCNN	274	10412	284	12466
CatBoost	398.7	3493.2	584.2	3203.4
RNN	-	-	-	-
Proposed	267.08	1575.88	225.53	3905.33

In addition to the CatBoost algorithm, our proposed method produces Scores values that are reduced by at least 1644.67 on the subset FD004, i.e., by 29.63%. This is the benefit of using a concurrent neural network, which extracts both temporal and spatial features of the signal. Although the performance of the proposed model on FD004 is not as good as the best results, taken together, the model performs very competitively on the subsets.

6. Conclusions

In this paper, through the study of intelligent prediction of aero-engine residual life (RUL), a model based on the concatenation of a one-dimensional CNN network and a BiLSTM network is proposed, aiming to improve the accuracy and reliability of the prediction. By utilising big data monitoring techniques and advanced deep learning algorithms, this study successfully addresses the limitations of traditional prediction methods in dealing with uncertainty in complex systems, and especially shows significant advantages in short remaining life prediction.

The research results in this paper show that the proposed concurrent neural network model can effectively extract and fuse the time series and spatial series features in aero-engine monitoring data, which improves the prediction performance of RUL. Empirical analyses show that the model achieves better results than the existing frontier methods on NASA's C-MAPSS dataset, especially in predicting the short remaining life, which significantly reduces the error, which is of great practical significance in identifying the potential failures in a timely manner and reducing the maintenance cost.

Future research can be carried out in depth in the following directions:

1. further optimise the model structure and explore more efficient feature extraction and fusion methods to adapt to different types of aero-engines and more complex operating conditions.
2. expanding the size and diversity of the dataset to include different types of aero-engines and more operating conditions to enhance the generalisation capability and robustness of the model.
3. incorporate domain expert knowledge to interpret and validate the model's predictions and improve the model's interpretability to better guide practical maintenance decisions.

Through the implementation of these recommendations, it is expected to further enhance the safety and affordability of aero-engines and make greater contributions to the development of China's aero-engine industry.

References

[1] Wu B, Zeng J, Shi H, Zhang X, Shi G, Qin Y. Multi-sensor information fusion-based prediction of remaining useful life of nonlinear Wiener process[J]. Meas Sci Technol. 2022;33(10):105106.

[2] Li N, Gebraeel N, Lei Y, Fang X, Caim X, Yan T. Remaining useful life prediction based on a multi-sensor data fusion model[J]. Reliab Eng Syst Saf 2021;208:107249.

[3] Cui L, Li W, Wang X, Zhao D, Wang H. Comprehensive remaining useful life prediction for rolling element bearings based on time-varying particle filtering[J].IEEE Trans Instrum Meas 2022;77:1–10.

[4] Li X, Jiang H, Liu Y, Wang T, Li Z. An integrated deep multiscale feature fusion network for aeroengine remaining useful life prediction with multisensor data[J].Knowl Based Syst 2022;235:107652.

[5] Liu J, Lei F, Pan C, Hu D, Zuo H. Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion. Reliab Eng Syst Saf 2021;214:107807.

[6] Guo R, Liu Z, Wei Y. Remaining useful life prediction for the air turbine starter based on empirical mode decomposition and relevance vector machine[J]. Transactions of the Institute of Measurement and Control. 2020;42(13):2578-2588.

[7] Chen, Jiaxian, et al.Aero-Engine Remaining Useful Life Prediction Method with Self-Adaptive Multimodal Data Fusion and Cluster-Ensemble Transfer Regression[J]. Reliability Engineering & System Safety, vol. 234, June 2023, p. 109151,

[8] Xia T, Song Y, Zheng Y, Pan E, Xi L. An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation[J]. Comput Ind 2020;115:103182.

[9] S. Xiang, Y.i. Qin, C. Zhu, Y. Wang, H. Chen. LSTM networks based on attention ordered neurons for gear remaining life prediction[J].ISA Trans., 106 (2020), pp. 343-354

[10] X. Liu, L. Liu, D. Liu, L. Wang, Q. Guo, X. Peng. A hybrid method of remaining useful life prediction for aircraft auxiliary power unit[J]. IEEE Sensors J., 20 (14) (2020), pp. 7848-7858.

[11] Zhang A, Wang H, Li S, Cui Y, Liu Z, Yang G, Hu J. Transfer learning with deep recurrent neural networks for remaining useful life estimation[J]. Appl Sci 2018;8:2416.