



Research on Failure Cause Analysis Method Based on Aircraft Maintenance Records

Fan Jiang^{1,2,3}(✉), Baohui Jia⁴, Jinglin Wang^{1,2,3}, and Guo Zheng^{1,2,3}

¹ Aviation Key Laboratory of Science and Technology on Fault Diagnosis and Health Management, Shanghai 201601, China

474994454@qq.com

² AVIC United Technology Center for Basic Research on Aircraft System Fault Diagnosis and Health Management, Shanghai 201601, China

³ AVIC Shanghai Aero Measurement and Controlling Research Institute, Shanghai 201601, China

⁴ School of Transportation Science and Engineering, Civil Aviation University of China, Tianjin 300300, China

Abstract. Due to the accumulation of a large amount of maintenance text information recorded by maintenance personnel based on experience in aircraft maintenance, due to its strong professionalism and diverse expression methods, intelligent analysis has not yet been achieved, making it difficult to fully utilize these data. In response to this issue, this article proposes an aircraft maintenance record analysis method based on the ALBERT-TextCNN model, which solves the fault causes to assist maintenance personnel in correct troubleshooting work. Firstly, the vectorization representation of the preprocessed text and the preliminary feature extraction are performed by using ALBERT; Secondly, the TextCNN model is used to extract deeper features to improve the accuracy of model analysis; Finally, compared with other commonly used text analysis models, the results show that the method used in this article can accurately analyze the causes of faults in aircraft maintenance records, effectively improving the accuracy and speed of analysis.

Keywords: Aircraft Maintenance Records · Fault Cause Analysis · ALBERT · TextCNN

1 Introduction

In recent years, aircraft flight conditions have become increasingly complex, and there are many causes of faults. Improper handling of any detail can have a significant impact on the quality of aircraft fault maintenance and handling, and lay safety hazards. The accuracy requirements for fault diagnosis of aircraft systems that may have abnormalities and faults are also increasing. The cause of a malfunction is the primary goal of fault diagnosis. Only after determining the cause of the malfunction can corresponding troubleshooting work be carried out to avoid greater economic losses and casualties in a timely manner [1]. Ground comprehensive diagnosis is an important component of aircraft health management. During the ground troubleshooting process, there is a

large amount of maintenance data, sometimes presented in the form of sensor collected signals, and sometimes in the form of descriptive text or maintenance records formed during manual maintenance. Therefore, how to analyze and extract useful and accurate information from a large amount of various forms of data is crucial for maintenance decision-making.

The maintenance text usually contains key features such as status information related to faults, methods used during troubleshooting, effectiveness of methods, and causes of faults [2], which can reflect information dimensions that sensor data does not possess. However, due to the complexity of maintenance records, they may not be directly used for the learning process of intelligent analysis. Therefore, few people analyze them and the data utilization rate is low, which seriously affects the accuracy of fault diagnosis for civil aircraft. If they can be fully utilized, they can be combined with sensor data to establish a connection between the observed aircraft appearance and the internal state of the aircraft, helping maintenance personnel quickly identify the cause of the fault and make correct maintenance decisions improves the level of fault repair while saving personnel time and cost, which is greatly beneficial for the fault diagnosis and analysis work of civil aircraft.

Effective diagnosis of aircraft system faults can be used to determine the operational status of the aircraft, facilitating its maintenance and operation, and is of great significance. With the development of information technology, intelligent technology is gradually being applied in the field of fault diagnosis. Chen et al. [3] developed a fault diagnosis method for engine gas path components that combines neural networks and D-S evidence theory to reduce misdiagnosis rates. Eroglu et al. [4] used convolutional neural networks to study the problem of actuator failures during the landing process of large passenger aircraft; Gururajan et al. [5] studied the fault diagnosis of airspeed tube of a small UAV using feedforward neural network.

At present, the research on fault diagnosis for structured data has achieved good results, but natural language processing can be used for the analysis of unstructured fault data. Reference [6] Two feature selection methods, serial chi square statistics and Dirichlet allocation, were proposed for log data of train control on-board equipment, which fused fault features from railway fault texts at both syntactic and semantic levels. A fault diagnosis model for on-board equipment was established based on support vector machines; Liu et al. [7] constructed a defect text classification model based on Convolutional Neural Networks according to the characteristics of power equipment defect texts; Zhao et al. [8] Based on fault text of railway on-board equipment, after feature extraction through LDA topic model, Bayes structure learning algorithm is applied to fault classification; Yan [9] et al. proposed an intelligent fault diagnosis method based on LSTM-SMOTE model for fault text number of metro engineering vehicles, which solved the problem of unbalanced fault categories.

Based on the above research analysis, although these research works have good application effects in multiple fields, there are still the following problems in the field of aviation maintenance data:

1. At present, there are few research results on fault analysis of aircraft maintenance records in the field of aircraft, and a large amount of aircraft maintenance text data has not been fully utilized;

2. The method of manually annotated information is still used for aircraft maintenance records, but with the sharp increase in data volume, the efficiency and scientificity of its analysis and diagnosis are difficult to guarantee, and it has a low level of intelligence;
3. The civil aircraft maintenance text contains a large number of professional vocabulary, and the application of traditional text analysis techniques is ineffective.

Therefore, this article establishes a fault cause analysis model for aircraft maintenance records based on the ALBERT TextCNN model to address the above issues. Firstly, the ALBERT [10] pre-training language model is established to represent the feature vector of the aircraft maintenance text at the sentence level to improve the representation ability of the text vector. Secondly, the output of the last Transformer layer is used as the input of the downstream TextCNN [11] model to maximize the Semantic information of different abstract levels and obtain the deep semantic features of each maintenance record. By mining the deep text features of maintenance record text content, the problem of low classification accuracy and high computational resource consumption in traditional classification algorithms for maintenance records can be solved. At the same time, it can help maintenance personnel locate faults as soon as possible, save maintenance personnel time and cost, and thus enhance the utilization value of aviation maintenance records in engineering.

2 Analysis of Aircraft Maintenance Record Text Data

The large amount of maintenance record data accumulated in the on-site work of aircraft maintenance is recorded in the database of maintenance records in natural language. The aircraft maintenance records (partial) are shown in Table 1. As textual data in the professional field, aircraft maintenance records have the following characteristics:

1. Aircraft maintenance records mainly contain multiple sets of information. The ATA chapter, fault description, and corrective measures are the main analysis content.
2. All are short texts, with a length of no more than 120 words, and there is a problem of semantic sparsity.
3. Due to the fact that aircraft maintenance is filled out by different personnel without unified standards, there is a problem of colloquialism and non-standardization.
4. There is noise data in the maintenance records, which has no effect on semantic feature extraction. For example: “&”, “。” et al.
5. The maintenance record field has strong characteristics, including a large number of proprietary names and equipment codes, and the commonly used word segmentation lexicon cannot fully cover it.

The structure of aircraft is very complex, and its faults have characteristics such as diverse types and complex classification. This article uses some historical maintenance records of a certain aircraft fleet, which includes 29562 maintenance records. The distribution of fault cause categories is shown in Fig. 1.

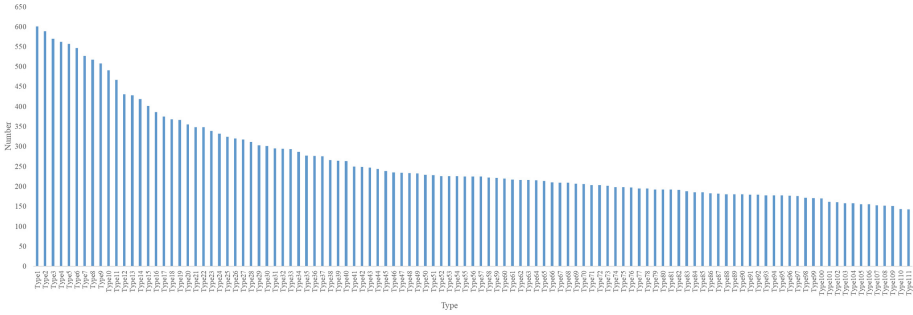


Fig. 1. Cause and quantity of faults

3 Analysis Model for Fault Causes of Aircraft Maintenance Records

3.1 Model Framework

The framework of the ALBERT-TextCNN fault cause analysis model for aircraft maintenance records proposed in this article is shown in Fig. 1. The framework is mainly divided into four layers (Fig. 2):

1. Repair record preprocessing layer
2. Input vector and feature presentation layer
3. Feature extraction and classification layer of aircraft maintenance records
4. Output results and model evaluation layer

3.2 Preprocessing of Aircraft Maintenance Record Datasets

The main preprocessing methods for the text characteristics of aircraft maintenance records analyzed above include:

1. Data deletion and integration. Delete the data columns unrelated to the analysis in this article: serial number, model, machine number, occurrence date, fault report, and maintenance level. Merge the ATA chapter, fault description, and corrective measures into one column for subsequent analysis.
2. Deactivate word filtering. For some noise words that are not related to the classification results, such as “在”, “的” and other common stop words, delete them. At the same time, remove words such as “发现”, “根据”, “同时” from the maintenance record case text as stop words.
3. Text cleaning. Characters such as punctuation mark and special symbols that do not carry valid text feature information or actual meaning may appear in maintenance records. To avoid affecting model training processes such as feature extraction, regular expressions need to be used to delete them.

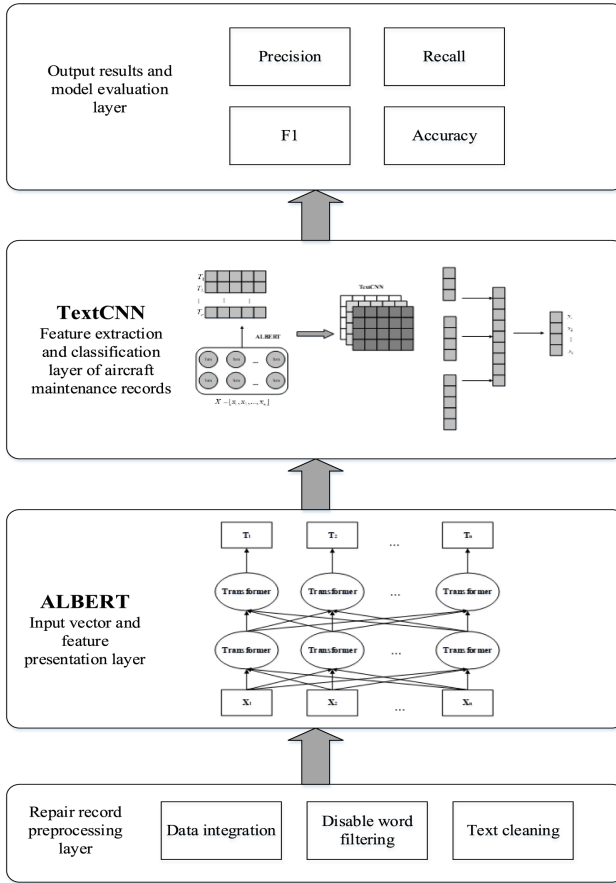


Fig. 2. Framework of aircraft maintenance record analysis model Based on ALBERT-TextCNN

3.3 Input Vector Representation of Aircraft Maintenance Records

First, the original input statement will be changed into a word sequence $w = (w_1, w_2, \dots, w_n)$ after passing through the input layer and enter into the presentation layer of the ALBERT input vector. This layer will first convert the word sequence into a vector sequence X of the input neural network, add special characters $[CLS]$ and $[SEP]$ at the beginning and end of the input text, express the content of the input text as serialized maintenance record text information through the input layer and the coding layer, and finally input it into the ALBERT network. As shown in Fig. 3.

$$\begin{cases} X = [x_1 x_2 \dots x_n] \in R^{n \times H} \\ x_i = E_i^t + E_i^p + E_i^s \quad i = 1, 2, \dots, n \end{cases} \quad (1)$$

In the formula, H is the vector dimension; E_i^t is the word embedding encoding of the word sequence; E_i^p is the encoding of positional information for word sequences;

E_i^s is the encoding of word sequence statement information, mapping the above three to high-dimensional addition to obtain the input sequence X of the ALBERT layer.

In order to improve the training speed of the model, ALBERT uses decomposition of the embedded word vector matrix to reduce the number of model parameters. Decompose the word vector matrix into two low dimensional word vector matrices and map them to the hidden layer. Suppose that the size of the glossary is V , the size of the word embedding layer is E , and the size of the hidden layer is H . The original dimension of the parameter matrix was $O(V \times H)$, and after ALBERT decomposes the embedded word vector matrix, the dimension of the parameter matrix becomes $O(V \times E + E \times H)$. When the hidden layer size H is much larger than E , the number of parameters in the model will be greatly reduced [12].

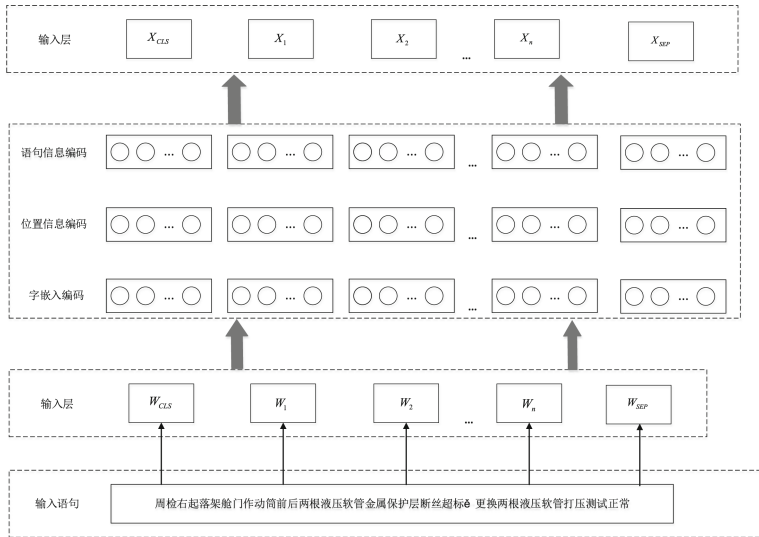


Fig. 3. Input vector representation of aircraft maintenance records

3.4 Aircraft Maintenance Record Feature Presentation Layer

The second step of the ALBERT model proposed in this paper is to input the text Initialization vector matrix X after vector expression into the Transformer [13] model loaded with pre training parameters. In the process of stack calculation, the output vector of the upper layer is used as the input vector of the lower layer, and the whole text sequence input is read in both directions at one time to fully learn the context information of words. Its structure is shown in Fig. 4.

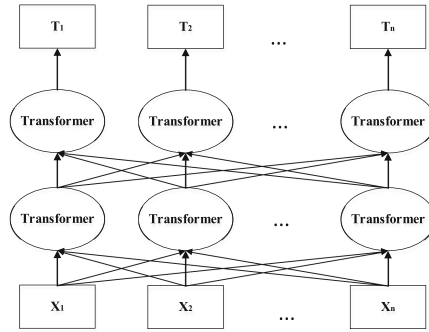


Fig. 4. ALBERT model structure

Among them, X_1, X_2, \dots, X_n represent the vector representations of words in the text sequence of aircraft maintenance records; T_1, T_2, \dots, T_n represent the feature vector representations of the obtained aircraft maintenance record.

The internal structure of the Transformer model is shown in Fig. 5.

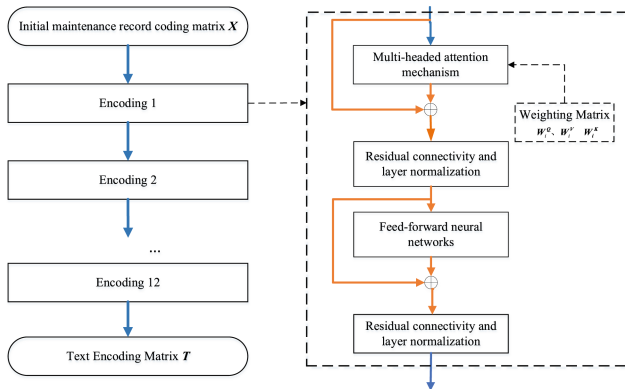


Fig. 5. Transformer model structure

The most important module in the Transformer structure is the multi-head attention mechanism, which can parallelly process encoded word vectors, automatically learn the contextual relationships between characters, and jointly focus on information from different presentation subspaces at different positions. Firstly, initialize multiple sets of Q, K, V matrixes randomly based on the input initial maintenance record coding matrix.

$$Q = XW^Q + b^Q \tag{2}$$

$$K = XW^K + b^K \tag{3}$$

$$V = XW^V + b^V \tag{4}$$

$$Attention(Q, K, V) = Soft \max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{5}$$

$$MultiHead = concat(head_1, head_2, \dots, head_n) \tag{6}$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{7}$$

In the formula, W^Q, W^K, W^V are the coefficient matrixes; b^Q, b^K, b^V are the bias vectors; $head_i$ represents the attention result of the i -th head, and W_i^Q, W_i^K, W_i^V are the coefficient matrixes of the i -th heads; d_k is the dimension of the Q matrix for each word; $concat(\bullet)$ is a concatenation function.

Residual connection and layer normalization are used to solve the common problem of feature extraction capacity saturation caused by model depth, resulting in gradient disappearance. The full connection layer is used for forward calculation of feedforward neural network, which provides the output of multi-head attention with the ability of nonlinear transformation and improves the representation ability of the model.

In the ALBERT model training process, in order to improve efficiency, a shared fully connected layer and an attention mechanism layer are used in the Transformer model, that is, all parameters of the shared hidden layer are shared to reduce the number of parameters and improve the model training speed. Due to the more complete semantics of multiple consecutive words, the model is trained by randomly covering multiple consecutive words, which increases the difficulty of pre-training and is beneficial for fault localization tasks in downstream aircraft maintenance text data.

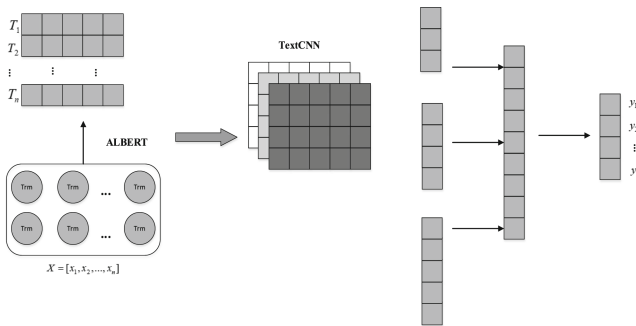


Fig. 6. Deep feature extraction structure for maintenance records

3.5 Feature Extraction and Classification of Aircraft Maintenance Records

In the above ALBERT training process, the method of sharing parameters across layers was used to improve the training speed, but at the same time, it also reduced the accuracy of the model to a certain extent. Therefore, for this problem, this paper uses the TextCNN model to conduct the convolution pooling operation on the feature matrix T of the

maintenance records extracted by the Transformer encoder through the use of sliding windows of different sizes, capture the local features of the text sequence for combination and filtering, and extract the text Semantic information at different abstract levels [14]. Further extract deeper features to obtain high-level feature vector representations of aircraft maintenance records, as shown in Fig. 6.

$$\mathbf{T} = [\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_n] \quad (8)$$

$$\mathbf{T}_{1:n} = \mathbf{T}_1 \oplus \mathbf{T}_2 \oplus \dots \oplus \mathbf{T}_n \quad (9)$$

Among them, \oplus is the connector; $\mathbf{T}_{1:n}$ represents concatenating sentences with a length of n into a matrix.

Feature \mathbf{c}_i through convolution operation is:

$$\mathbf{c}_i = f(\mathbf{w} \bullet \mathbf{T}_{i:i+h-1} + \mathbf{b}) \quad (10)$$

$$\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{n-h+1}]$$

where \mathbf{b} is the offset vector, $f()$ is the nonlinear activation function, and \mathbf{w} is the convolution kernel; \mathbf{c} is the feature vector obtained after the convolution operation.

After extracting local features, the most representative feature vector, namely the global feature vector, is extracted using the maximum pooling method.

$$k = \max\{c\} \quad (11)$$

Finally, enter the fully connected layer for fault localization.

$$y = \text{soft max}(k) \quad (12)$$

Among them, y represents the cause of the fault recorded in the aircraft maintenance record.

4 Experimental Results and Analysis

4.1 Experimental Environment and Data

The experimental environment in this paper is as follows: Windows 11; AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz; Python 3.7; tensorflow 1.15.

Firstly, the aircraft taken in this paper is subjected to data preprocessing operation, and the preprocessed aircraft maintenance records are shown in Table 2. Secondly, they are divided into training set, test set and validation set according to the ratio of 8:1:1 (data volume: 23650, 2956, 2956 respectively), and finally input into ALBERT-TextCNN model for failure cause analysis.

Table 2. Aircraft maintenance record pre-processing results (example)

Before pre-treatment	After pre-treatment
<p>“The #1 main wheel is wear. Due to the short stop troubleshooting (i.e., replacing the #1 main wheel) time is not enough and this fault (i.e., this #1 main wheel) meets the criteria for a short stop release, so release the aircraft and wait for the post-flight replacement. POST FLIGHT CHECK THE NO. 1 MAIN WHEEL HAS OVER THE LIMIT. Post-flight replacement of the #1 main wheel, the check is normal!” 4</p>	<p>No. 1 main wheel is wear. Short stop troubleshooting (Replacement of No. 1 Main Wheel) Insufficient time fault (No. 1 Main Wheel) meets the short stop release criteria, released, replaced after flight. Post flight check the no.1 main wheel has over the limit</p>

4.2 Analysis of Experimental Results

The aircraft maintenance records were analyzed by ALBERT-TextCNN model and the results were obtained as Table 3.

Table 3. Failure analysis results

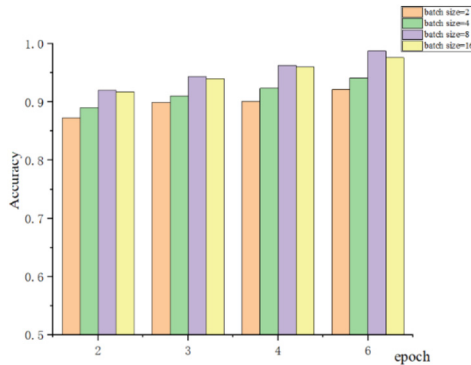
Serial number	Fault description	Cause of Failure
1	Post-cruise inspection APU post-cruise report has IGV fault message, replace APU IGV actuator cylinder check normal	IGV actuator
2	Troubleshooting. In order to verify the BXXXX failure, the primary heat exchanger of the left module was strung with the heat exchanger of the left module of the BXXXX, and the test was normal. PN:XXX-XXX	Heat exchanger
3	After the flight, found that the magnetic plug of the left main oil return filter has a piece of 1 × 3 MM metal chip. Replace the left main oil return filter and oil supply filter, test check, there is no metal chip in 7 magnetic plugs, and the oil supply and return filters are normal	Magnetic plug
4	The unit reflected that the brake is not sensitive, replace the two-way valve, the operation check is normal	Brake valve
5	Rattling of the front wheels during turns. The front wheels and the front landing gear turning mechanism were inspected and no damage was found, and the front wheel turning mechanism was lubricated. Check normal	Front wheel turn
...

Table 4. ALBERT-TextCNN model training parameters

Training parameters	Parameter value
max_seq_length	128
learning_rate	3e-5
dropout (TextCNN)	0.1
kernel size	3, 4, 5
filter	128
Layer	1

The training parameters of the ALBERT-TextCNN model are shown in Table 4 above.

To study the influence of model-related parameters on the results of fault cause analysis of aircraft maintenance text, this paper first investigates the influence of epoch and batch size on the accuracy of the model, and the results are shown in Fig. 7. From the results, we can see that the accuracy of the model is highest when epoch = 5 and batch size = 8, because too small batch size will make the model difficult to converge and too large to fall into local optimum.

**Fig. 7.** Impact of epoch and batch size on accuracy

Next, the effect of the number of attention heads on the accuracy of the model is investigated, and the results are shown in Fig. 8. From the figure, it can be seen that the accuracy rate is highest when the number of attention heads is 12. Increasing the number of attention heads can enhance the feature extraction ability of the maintenance records, but too many of them will cause redundancy in the information captured between the attention heads, thus reducing the accuracy rate.

To investigate the effectiveness and superiority of the ALBERT-TextCNN method proposed in this paper in the analysis of the causes of aircraft maintenance record failures, TextCNN, BiLSTM, BERT and ALBERT models were selected as controls. Their training losses are shown in Fig. 9.

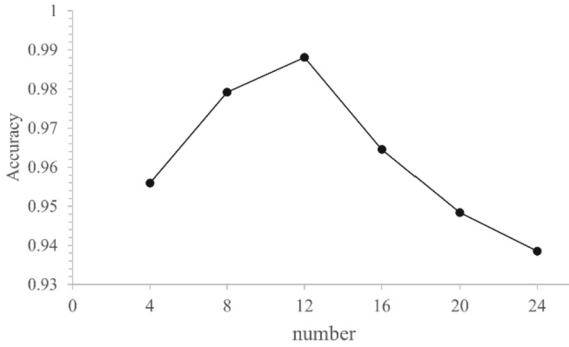


Fig. 8. Effect of number of attention heads on accuracy

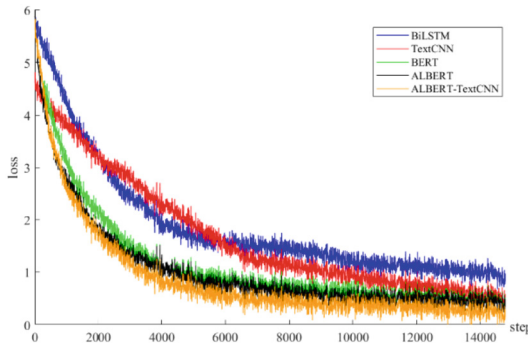


Fig. 9. Model training loss

Figure 9 shows that the training losses of the ALBERT-TextCNN model used in this paper outperforms other models to a certain extent and shows more stable convergence (Table 5).

Table 5. Experimental results for each model

Model	Accuracy/%	Precision/%	Recall/%	F1/%
TextCNN	70.81	72.56	71.13	71.84
BiLSTM	87.94	88.43	89.57	88.50
BERT	98.79	96.03	96.62	96.32
ALBERT	96.83	94.75	93.91	94.33
ALBERT-TextCNN	98.91	97.37	96.89	97.13

The ALBERT-TextCNN model solves the problem that the TextCNN model ignores the local information of text during training, and it uses the Transformer structure to fuse the global semantic information, which solves the distance-dependent problem of

the BiLSTM model. Problem. However, compared with the BERT model, the accuracy rate and F1 value are not significantly improved, but its training time and processing speed are significantly better than the BERT model, as shown in Table 6. Therefore, the ALBERT-TextCNN model can guarantee a high diagnosis rate while having a relatively fast diagnosis speed, which improves the engineering practicality of the model.

Table 6. Model training time comparison

Model	Training time/h
BERT	13
ALBERT	6.5
ALBERT-TextCNN	7

5 Conclusion

For the current aircraft maintenance records still rely on manual analysis, and the analysis has the problem of low efficiency, this paper combines the characteristics of aircraft maintenance text data, and proposes a maintenance text fault analysis method based on ALBERT-TextCNN model. Firstly, we analyze the characteristics of maintenance records and use them to choose to build an ALBERT model to extract text features by word, and at the same time, we combine the shortcomings of ALBERT model to build a TextCNN model to extract text semantic information at different abstraction levels. The results show that the method can effectively analyze the causes of faults in aircraft maintenance records, and performs better than other commonly used models, and has a faster analysis speed while ensuring a higher accuracy rate.

References

1. Wang, R.G., Wu, J., Liu, C., Yang, H.Y.: Method for identifying the causes of aircraft equipment failures based on maintenance logs. *J. Softw. Sci.* **30**(5), 1375–1385 (2019)
2. Du, X.M., Qin, J.F., Guo, S.Y., Yan, D.F.: Text mining of typical fault cases of power equipment. *High Volt. Technol.* **44**(04), 1078–1084 (2018)
3. Chen, T., Sun, J.G., Hao, Y.: Aero-engine airway fault diagnosis based on neural network and evidence fusion theory. *J. Aeronaut.* **27**(6), 1014–1017 (2006)
4. Eroglu, B., Sahin, M.C., Ure, N.K.: Autoland control system design with deep learning based fault estimation. *Aerosp. Sci. Technol.*, 102 (2020)
5. Gururajan, S., Fravolini, M.L., Chao, H., et al.: Performance evaluation of NN based approaches for airspeed sensor failure accommodation on a small UAV. In: *Mediterranean Conference on Control and Automation*. IEEE (2013)
6. Wang, F., Xu, T., Tang, T., et al.: Bilevel feature extraction based text mining for fault diagnosis of railway systems. *IEEE Trans. Intell. Transp. Syst.* **18**(1), 49–58 (2017)
7. Liu, Z.Q., Wang, H.F., Cao, J., et al.: Research on power equipment defect text classification model based on convolutional neural networks. *Power Grid Technol.* **42**(02), 644–651 (2018)

8. Zhao, Y., Xu, T.H.: Fault diagnosis of on-board equipment of high-speed railway signal system based on text mining. *J. Railw.* **37**(08), 53–59 (2015)
9. Yan, S., Xu, Y.N., He, W.T.: Research on intelligent fault diagnosis of metro engineering vehicles based on natural language processing. *J. Weapon Equip. Eng.* **43**(09), 101–108 (2022)
10. Lan, Z., Chen, M., Goodman, S., et al.: ALBERT: a Lite BERT for self-supervised learning of language representations (2019)
11. Mikolov, T., Chen, K., Corrado, G., et al.: Efficient estimation of word representations in vector space. *Comput. Sci.* (2013)
12. Zhang, D.Z., Fan, X.X., Xie, Y.H., et al.: A traditional Chinese medicine visceral localization model based on ALBERT and bidirectional GRU. *J. Eng. Sci.* **43**(9), 1182–1189 (2021)
13. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. In: *Annual Conference on Neural Information Processing Systems 2017*, pp. 5998–6008. MIT Press, Massachusetts (2017)
14. Kim, Y.: Convolutional neural networks for sentence classification. Eprint arXiv (2014)