Application and Study of Artificial Intelligence in Railway Signal Interlocking Fault

Hongwei Liang[1]*, Xiuxuan Wang[1], Anjali Sharma[2], Mohd Asif Shah[3]

[1]Zhengzhou Railway Vocational & Technical College, Zhengzhou, Henan ,451460, China

[2]School of Biological and Environmental Sciences, Shoolini University of Biotechnology and Management Sciences, Solan 173229 (H.P.), India

[3]Bakhtar University, Kabul, Afganistan

Emails: hongweiliang7@163.com, xiuxuanwang9@126.com, anjalisharmaas8749347@gmail.com, ohaasif@bakhtar.edu.af

The rapid development of railway transportation towards high speed, high density and heavy load has led to even higher requirements for the safety of railway signal equipment. The safety of railway signal equipment is an important part of ensuring railway traffic safety, thus, it is very necessary to study a system that can diagnose the fault of railway signal equipment according to the actual situation. This article utilizes the deep learning algorithm of artificial intelligence for investigating the interlocking faults in the railway transportation. This paper uses ADASYN data synthesis method to synthesize few category samples, uses TF-IDF to extract features and transform vectors, and proposes a deep learning integration method based on combined weight. The results show that BiGRU has better overall classification performance when evaluated on the index of primary and secondary fault classification accuracy. The classification accuracy improvement of 5% is achieved for primary fault classification and the comprehensive evaluation index of secondary fault classification is improved by about 9%. It was revealed that when compared with ADASYN + BiLSTM neural network, the comprehensive evaluation index of primary fault classification accuracy is improved by about 6%, and the comprehensive evaluation index of secondary fault classification is improved by about 10%. It is demonstrated that deep learning integration is an effective method to improve the classification performance of turnout fault diagnosis model.

Povzetek: Za železniški sistem je bila uporabljena metodologija globokih nevronskih mrež za iskanje napak v signalih.

1 Introduction

With the gradual increase of railway traffic density and operation speed in China, it is difficult to avoid various faults of railway signal equipment. If the faults cannot be handled in a short time, they will have a great impact on traffic safety, and even lead to the hidden dangers of major accidents, so as to reduce the efficiency and safety of railway operation. At the same time, it also brings new challenges to railway signal equipment maintenance personnel to check and maintain signal equipment timely and accurately.

High speed railway signal equipment is an important infrastructure to ensure high-speed train operation. The maintenance quality of signal equipment directly affects the traffic safety and transportation efficiency of high-speed railway. Signal equipment fault is diagnosed and handled according to the experience and knowledge of on-site maintenance personnel, which is easy to cause maintenance judgment error and maintenance time delay, and in serious cases, it will lead to equipment fault driving accident. The fault data of high-speed railway signal equipment records the fault phenomenon when the fault occurs in the form of text.

The fault phenomenon is analyzed based on text data mining technology. Combined with the diagnosis results of experts on the fault phenomenon, the fault diagnosis model of signal equipment is studied to assist maintenance personnel to quickly locate the fault location and cause according to the fault phenomenon. It will be of great significance to further improve the safety guarantee level of high-speed railway. The basic activity diagram of train fault detection method is shown in Figure 1.
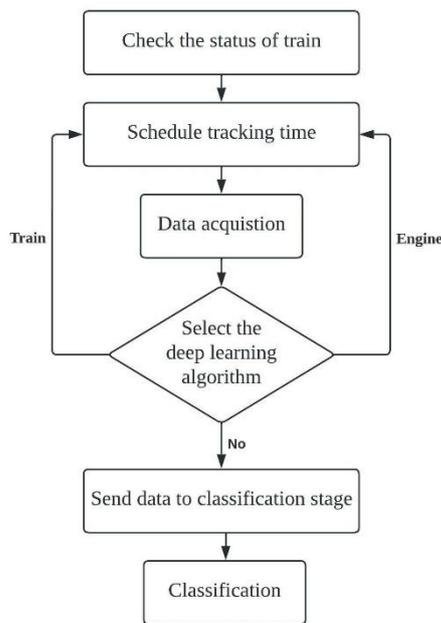
Figure 1: Activity diagram of Railway fault detection method

This limitation of imbalanced faults of different signal equipment is addressed in this article. In order to study the signal equipment fault diagnosis method based on unbalanced samples based on text mining technology, two problems need to be solved: one is the processing of unbalanced samples, and the other is the construction of fault diagnosis and classification model.

This article contributed in mainly using two methods to solve the sample imbalance problem: one is to synthesize the sample data by using data enhancement, under sampling or oversampling, and data generation methods such as SMOTE (Synthetic Minority Oversampling Technology) and ADASYN (Adaptive Synthetic Sampling). The other is to adjust the parameters of different categories for the classification learning algorithm. The sample synthesis algorithm can appropriately synthesize a few categories of samples according to the distribution of the overall samples, and can ensure that the sample data is not repeated. There are several articles which uses SVM-SMOTE method to automatically synthesize the few category samples of signal equipment fault, so as to solve the problem of signal equipment fault sample imbalance. This article utilizes the deep learning algorithm of artificial intelligence for investigating the interlocking faults in the railway transportation. This paper uses ADASYN data synthesis method to synthesize few category samples, uses TF-IDF to extract features and transform vectors, and proposes a deep learning integration method based on combined weight. The outcomes obtained for the proposed method reveals that BiGRU has better overall classification performance when evaluated on the index of primary and secondary fault classification accuracy.

The rest of this article is structured as: review of literature is provided in section 2 followed by research methodology involved in analysis of fault diagnosis of railway unlocking system in section 3. Section 4 provides the experimental results and discussion along with concluding remarks in section 5.

2 Related work

In this section various state-of-the-art work in the field of railway signal interlocking fault based on artificial intelligence and other technologies is presented.

With the advent of the intelligent era, artificial intelligence has become the mainstream technology in the world, and artificial intelligence technology has laid a solid research foundation [1]. Paek and Kim explores the future direction of education by examining the current impact of artificial intelligence and predicting the future impact [2]. Interlocking is a railway system, which can automatically control safety management route change and avoid train collision and derailment. Dobias and Kubatova analyzes the latest technologies used in several commercial interlocking equipment, and proposed the design and implementation of an interlocking system architecture based on FPGA technology [3]. In order to solve the problem of channel estimation based on demodulated reference signal (DMRs) in railway tunnel scene, Skiribou *et al.* proposed a deterministic model to accurately generate time-varying channel response [4]. Kiedrowski and Saganowski introduced a scheme of applying PLC technology to railway light signs. This paper introduces the structure of the network and a group of equipment to realize this specific type of wired sensor network, which is used to monitor the railway led sign network and maintenance parameters [5]. Yang *et al.* analyzed the requirements of clock synchronization of signal ground equipment in combination with the application status of clock synchronization of ground equipment in high-speed railway signal system. By analyzing the advantages and disadvantages of the world's mainstream satellite navigation system and the requirements of China's railway signal system, Beidou time service technology is selected as the clock synchronization technology of the ground equipment of high-speed railway signal system, and the overall scheme based on Beidou time service technology is constructed [6]. In order to evaluate the network access performance of railway signal equipment machine communication (MTC) in the next generation intelligent transportation system, Lin *et al.* divided the railway signal equipment machine communication traffic prediction model into station indoor model, station outdoor model and station outdoor model, and calculated the traffic and signaling overhead of the three models respectively. Based on Poisson distribution and Markov renewal process, an improved Markov modulated poisson process (immpp) for source traffic model is designed [7]. Wang *et al.* combined with the new technical characteristics of high-speed railway, analyzed the current situation of lightning protection technology and lightning faults of foreign railway signal equipment. At the same time, the functions of intelligent technologies such as lightning activity location and lightning fault diagnosis are

introduced, and the development direction of railway lightning protection in the future is prospected according to the characteristics of this technology [8]. In order to realize the real-time acquisition, monitoring and management of the technical status of railway signal equipment and meet the multi-dimensional business needs of railway signal system information sharing, data mining, analysis and display, Sahal *et al.* put forward the national technical big data platform of railway signal equipment on the basis of analyzing the current situation of railway signal system and the significance of signal big data platform construction [9]. Based on the common signal system equipment of rail transit stations at home and abroad, Cao *et al.* analyzed the common faults and their settings of the system, studied the common faults analysis, design and construction of the signal system, and developed the railway signal fault setting training system based on the core concept of fault safety design [10]. In order to solve the problem of railway transportation safety, Dong *et al.* carried out detection experiments on simulated images and real videos of railway signal lights based on machine vision. The image features of railway signal lights in different color spaces and their influence on railway signal light recognition are discussed [11].

Railway signal equipment safety is an important part of ensuring railway traffic safety, thus, it is very necessary to study a system that can diagnose the fault of railway signal equipment according to the actual situation. The literature suggests that there are many studies on using data synthesis method to solve the sample imbalance based on the deep learning of artificial intelligence approach [12-15]. This paper diagnoses the fault of high-speed railway signal equipment, improves the performance of equipment fault diagnosis, so as to improve the safety of railway.

3   Research methods

This section includes the description of small category sample generation based on ADASYN. The fault text features of high-speed railway signal are also represented in this section and fault diagnosis model is presented.

High speed railway signal fault diagnosis forms a turnout fault diagnosis model with deliverable evaluation indexes through the training and optimization of the fault diagnosis model based on deep learning integration [16]. The turnout fault phenomenon of high-speed railway is input into the fault diagnosis model, and the model automatically outputs the type and cause of the fault, so as to realize the intelligent diagnosis of turnout equipment fault [17-19]. The architecture of this research work is depicted in Figure 2.
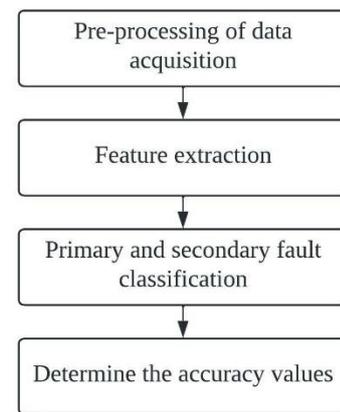

Figure 2: Architecture of research work

The basic structure of this research work includes pre-processing of data acquired from various sources. Further, the feature set is extracted followed by the classification of primary and secondary faults [20, 21]. At the final stage, accuracy values are determined for the proposed architecture. The development of artificial intelligence and Internet of Things is considered for several industrial applications and contributing towards social life [22-25].

3.1   Small category sample generation based on ADASYN

ADASYN adaptive synthesis oversampling method is to adaptively synthesize a small number of samples according to the distribution of a small number of samples, and synthesize fewer samples where it is easy to classify and more samples where it is difficult to classify. The key of the synthesis algorithm is to find a probability distribution r_i. Put r_i is the criterion for determining how many samples should be synthesized for each small category sample.

The proportion of the number of secondary categories included in each primary fault category of high-speed railway signal turnout fault is 12:17:8:11:7:1:7. Therefore, ADASYN is used to synthesize fewer secondary fault category samples, and the imbalance of primary fault categories can be solved at the same time. The process of using ADASYN to adaptively generate turnout secondary few category samples is as follows:

Step 1: Calculate the unbalance degree of few categories, $d = m_s/m_l$, m_s and m_l represent the number of samples with few categories and multiple categories respectively, $d \in (0, 1]$.

Step 2: Calculate the total number of small category samples to be synthesized, $G = (m_l - m_s) \times \beta$, $\beta \in (0, 1]$, indicating the expected imbalance degree of the whole sample after adding the synthetic sample, $\beta = 1$ means that the sample category is completely balanced after adding the synthetic sample.

Step 3: For each sample of a few categories x_i. Find their K-nearest neighbors in *n*-dimensional space and

calculate the ratio $r_i = \Delta_i/K(i = 1,2,...,m)$, m is the total number of samples, Δ_i is the number of multiclass samples in the k-nearest neighbor of x_i, so $r_i \in (0, 1]$。

Step 4: According to $\hat{r}_i = r_i/\sum_{i=1}^{m_s} r_i$, regularize r_i. So r_i is the probability distribution, and $\sum \hat{r}_i = 1$.

Step 5: Calculate the number of samples $g_i = \hat{r}_i \times G$ to be synthesized for each small category sample x_i, G is the total number of synthetic samples.

Step 6: According to the above steps, calculate the number of samples g_i synthesized by each small category sample x_i.

3.2 Fault text feature representation of high-speed railway signal equipment

TF-IDF is a text feature representation method based on weighting idea. Its core idea is that if a word appears frequently in one document and low in other documents, it indicates that the word has high recognition in the document and assigns its high weight. The feature extraction of signal equipment fault text first needs to realize Chinese word segmentation [26-29]. Because the high-speed railway signal equipment fault text data contains professional words such as switch machine, red light band and sealer, this paper constructs railway signal professional thesaurus and loads the thesaurus into Jieba word segmentation tool to realize the accurate word segmentation of fault text.

Text frequency (TF) in TF-IDF refers to the frequency of a given word in the document. For a given word t_i. In a document d_j, the degree of importance can be expressed as:

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{1}$$

Where: $n_{i,j}$ is the number of occurrences of the i-th word in document d_j. $\sum_k n_{k,j}$ is the total number of occurrences of each word in document d_j.

The inverse file frequency IDF is a measure of the general importance of a word. Its calculation formula is as follows. The larger the IDF, the better the ability to distinguish categories.

$$IDF_i = \log_2 \frac{|D|}{1 + |j : t_i \in d_j|} \tag{2}$$

Where: D is the total number of sample files, $|j : t_i \in d_j|$ contains the number of documents in the word. If the word is not in the sample, it will cause the denominator to be zero. Therefore, adding 1 to the denominator is to avoid the situation that the denominator is 0.

$W_{i,j} = TF_{i,j} \times IDF_i$. Weight $\omega_{i,j}$ of words is obtained by multiplying the word frequency in the document by the low file frequency of the word in the whole document set.

According to the TF-IDF feature weight calculation method, the characteristics of turnout fault samples based on text are calculated. The characteristics of a turnout fault sample are expressed as $d_i = [\omega_i^1 \ \omega_i^2 \ ... \ \omega_i^m]$, m is the length of the sample, and the primary fault category and secondary fault category are expressed as matrix $\mathbf{y}_1$ and $\mathbf{y}_2$ by one hot coding vectorization, $y_i = [0 \ 1 \ 0 \ ... \ c - 1]$, c is the total number of categories, and the fault level I category feature is expressed as $D_{L1} = [d_i \ y_1]$, ($i=1,2,..., n$), n represents the total number of samples. The label of fault level I is also input into the feature vector by Fault secondary feature as a feature, $D_{L2} = [[d_i \ y_1] \ y_2]$.

3.3 Deep learning integrated fault diagnosis model

Integrated learning is to combine multiple weak supervised learning models to get a better and more comprehensive supervised learning model. The high-speed railway turnout fault diagnosis model adopts BiGRU and BiLSTM neural networks as the weak supervised learning model, inputs the feature vectors extracted from the features into the embedded layer of BiGRU and BiLSTM neural networks respectively, and the two neural networks output the classification and prediction probability of the feature vectors in the Softmax layer through learning. The prediction results of the two neural networks are integrated and calculated by the combined weighted integration method, and finally the classification results of the input data by the deep learning integration model are output [30].

GRU and LSTM are variants of RNN neural network. Gating units are designed in neurons to effectively calculate and control the input and output of information, as shown in Figure 3. The design of this gating unit solves the problem of text sequence length dependence. Since the output of sigmoid function is 0 ~ 1, 1 can mean that the information is retained, and 0 means that the information is discarded, GRU and LSTM process the input information through sigmoid function, and tanh function processes the output information.
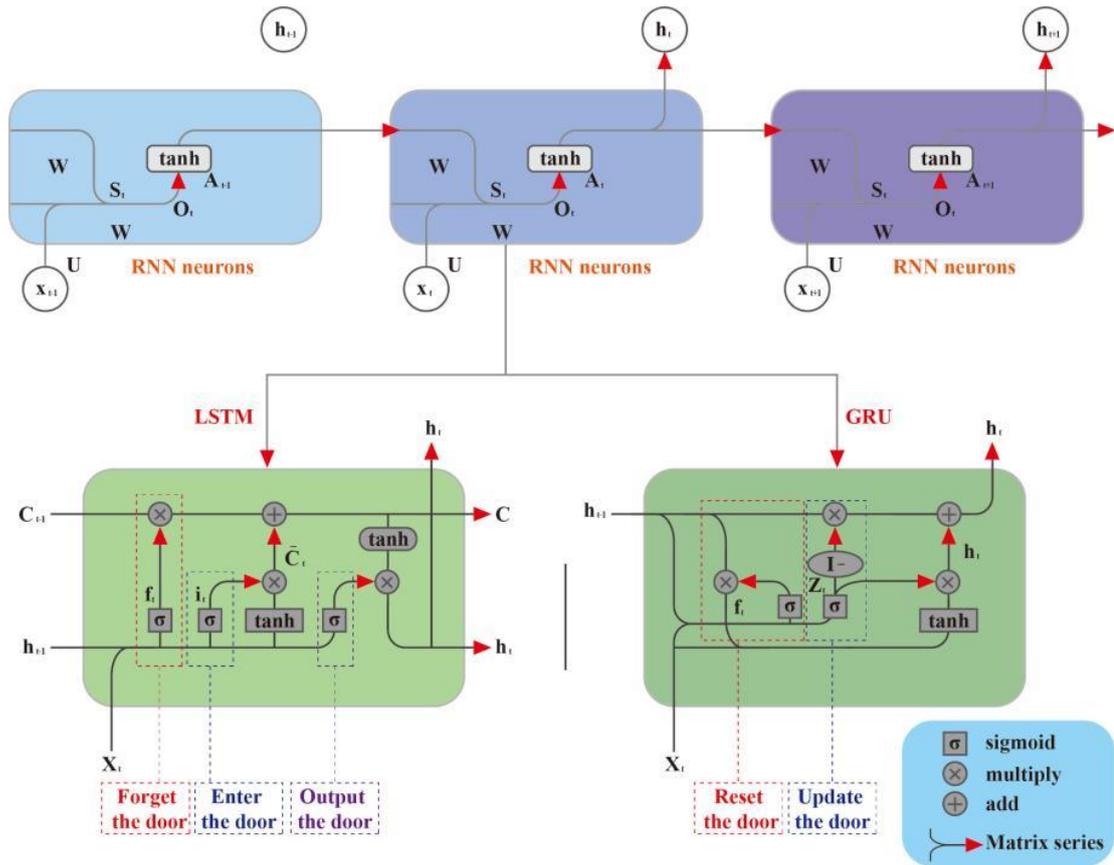
Figure 3: Structural units of RNN and its variant neurons

LSTM neural unit is composed of three gates, namely forgetting gate, input gate and output gate, as shown in Figure 3. LSTM first determines which information needs to be discarded through the forgetting gate, and calculates h_{t-1} x_i and output a vector between 0 and 1, the vector represents what information neuron C_{t-1} retains or discards. Then, the input gate is used to determine which information needs to be added in the neuron, and the candidate neuron $\widetilde{C}_t$ is obtained by tanh's calculation using h_{t-1} and x_i, which can be updated into the neuron. Finally, the output information is controlled by the output gate, and the LSTM neuron output is finally obtained by multiplying the 0 ~ 1 vector obtained by the output layer o_t and the neuron through the tanh layer.

$$f_t = \sigma(W_f \cdot [h_{t-1} \ \ x_i] + b_f) \tag{3}$$

$$i_t = \sigma(W_i \cdot [h_{t-1} \ \ x_i] + b_i) \tag{4}$$

$$\widetilde{C}_t = tanh(W_c \cdot [h_{t-1} \ \ x_i] + b_c) \tag{5}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{6}$$

$$o_t = \sigma(W_o \cdot [h_{t-1} \ \ x_i] + b_o) \tag{7}$$

$$h_t = o_t * \tanh(C_t) \tag{8}$$

where: * is Hadamard product operator, which means multiplication of elements at the same position of the matrix.

GRU is a variant of LSTM, as shown in Figure 3. It combines the forgetting gate and input gate into an update gate z_t. z_t controls how much information needs to be forgotten from the previous hidden layer h_{t-1}, how much information needs to be added to the current hidden layer $\widetilde{h}_t$, and then obtains h_t. Reset gate r_t controls how much previous information needs to be retained. When r_t is 0, $\widetilde{h}_t$ only contains the information of the current word.

$$z_t = \sigma(W_z \cdot [h_{t-1} \ \ x_t]) \tag{9}$$

$$r_t = \sigma(W_r \cdot [h_{t-1} \ \ x_t]) \tag{10}$$

$$r_t = \tanh(W \cdot [r_t * h_{t-1} \ \ x_t]) \tag{11}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \widetilde{h}_t \tag{12}$$

The combination weighted integration method of LSTM and GRU combines the overall classification performance of a single neural network with the classification performance of each category by assigning weights. The combination weighted integration method includes overall weight and category weight. The higher

the overall classification performance of a single neural network, the higher the overall weight will be allocated. According to formula (13) and formula (14), the lower the error proportion of neural network in category classification, the better classification performance it has in this category, the higher the category weight will be allocated. Then add the overall weight of the neural network and the category weight according to equation (15) to recalculate the predicted value of the neural network in each category. This combined weighted integration method can avoid the influence of few values and extreme values in the integration method.

$$\epsilon_{ij} = \frac{error\ Num_{ij}}{text\ Num_{ij}} \tag{13}$$

$$\alpha_{ij} = \begin{cases} \ln\left(\frac{1-\epsilon_{ij}}{\epsilon_{ij}}\right) & \epsilon_{ij} < 0.5 \\ 0 & \epsilon_{ij} \geq 0.5 \end{cases} \tag{14}$$

$$P_i = \sum_{j=1}^{n} (\omega_j + \alpha_{ij}) \cdot P_{ij} \tag{15}$$

Where: ϵ_{ij} is the classification error ratio of neural network j in category i. $text\ Num_{ij}$ is the total number of samples of category i; $error\ Num_{ij}$ is the number of classification error samples of neural network j in category i. α_{ij} is the category weight of neural network j in category i., ω_j is the overall weight of neural network j, and $\sum_{j=1}^{n} \omega_j$.

In order to improve the generalization ability of deep learning integration model, K-fold cross validation training model is adopted. K-fold cross validation is to randomly divide the whole training sample into K parts, one of which is used as the validation set and the other K-1 is used as the training set, and cycle K times until all data are selected once.

4 Results and Analysis

This section illustrates the result and analysis of overall weight distribution, weight calculation and the classification of deep learning integration model.

4.1 Overall weight distribution of BiGRU and BiLSTM

BiGRU and BiLSTM have the same network parameters, in which the embedded layer dimension is 100, the hidden layer dimension is 512, K-fold cross validation K = 4, the number of iterations is 50, and the batch size is 256. After TF-IDF feature extraction and vector representation, the training set and verification set synthesized by ADASYN are input into BiGRU and BiLSTM networks for training. The change of *loss* function value in the training process of the two neural networks is shown in Figure 4. It can be seen from Figure 4 that with the increase of iteration times, the *loss* value of BiGRU is lower than that of BiLSTM, indicating that its overall classification performance is

better. In the two neural networks, the *loss* function value of the primary classification is lower than that of secondary classification, indicating that the evaluation index of the primary classification of the neural network is higher than that of the secondary classification. Both neural networks are between 40 ~ 50 iteration rounds, and the *loss* function value tends to be stable, indicating that the number of iteration rounds of 50 can make the neural network training reach the best state.
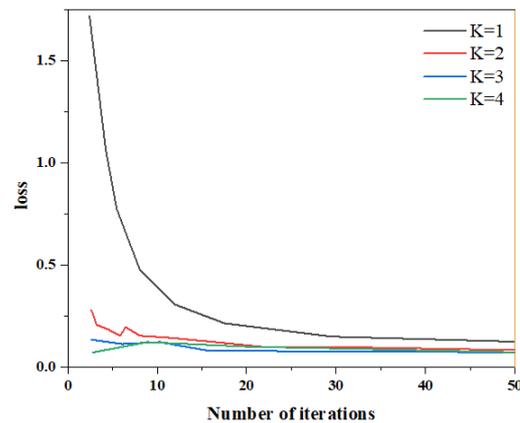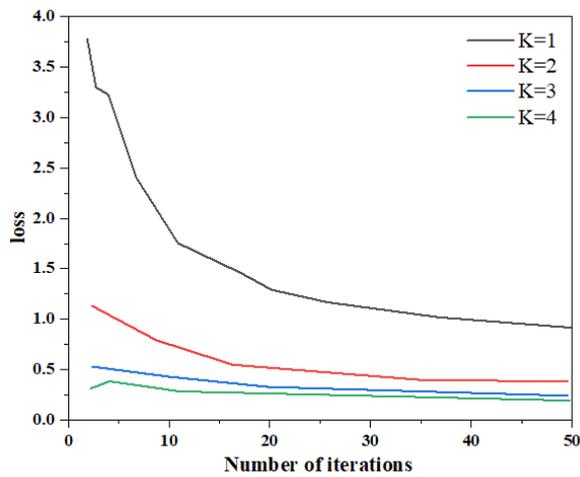

(a): BiGRU primary classification training process
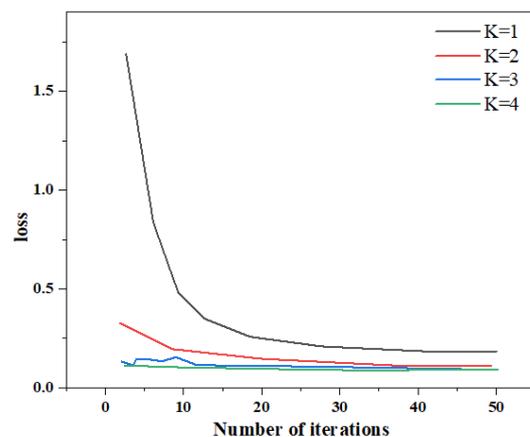
(b): BiGRU primary classification training process
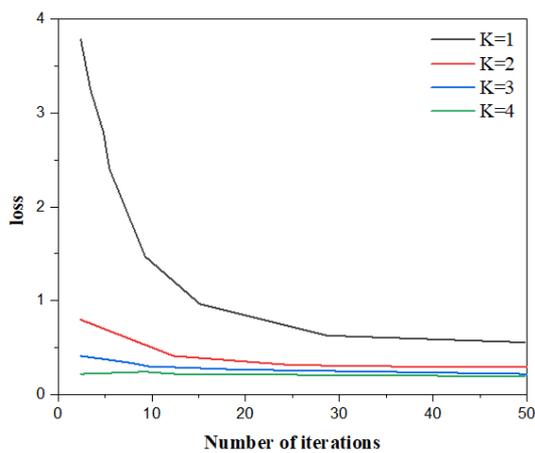
(c): BiGRU secondary classification training process

(d): BiGRU primary classification training process

Figure 4 (a, b, c, d): Variation of loss value in K-cross training of BiGRU and BiLSTM neural networks

After K = 4 training, 30% real samples are used to evaluate BiGRU and BiLSTM training models. The evaluation results are shown in Table 1 and is graphically presented in Figure 5.

Method	Level	Accuracy rate	Recall rate	F1 value
ADASYN ＋BiGRU	Primary fault classification	0.8742	0.8814	0.8779
	Secondary fault classification	0.7828	0.7421	0.7619
ADASYN ＋BiLSTM	Primary fault classification	0.8613	0.8765	0.8688
	Secondary fault classification	0.7601	0.7581	0.7591
BiGRU	Primary fault classification	0.7317	0.7098	0.7206
	Secondary fault classification	0.7081	0.6712	0.6891
BiLSTM	Primary fault classification	0.6912	0.7129	0.7019
	Secondary fault classification	0.6371	0.6214	0.6292

Table 1: Test results of K-fold cross validation + BiGRU and BiLSTM neural network
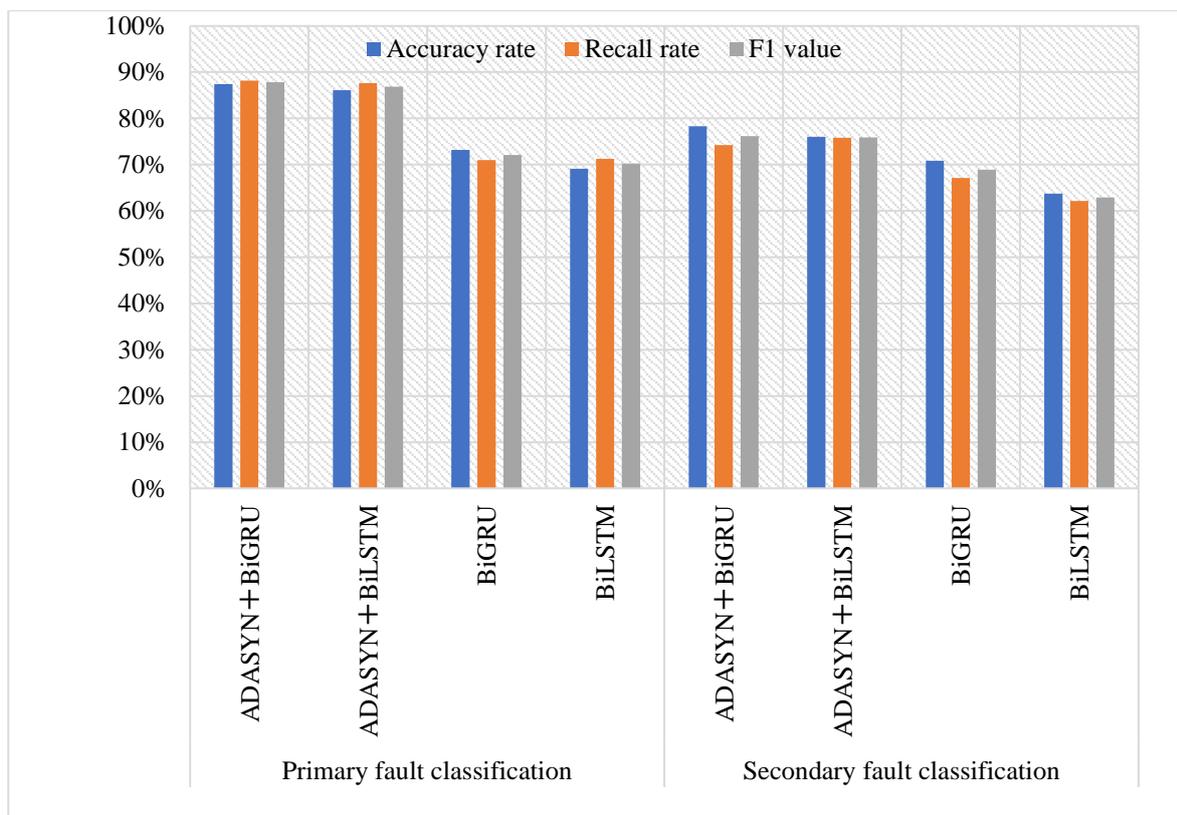
Figure 5: Graphical results of K-fold cross validation + BiGRU and BiLSTM neural network

It can be seen from Table 1 that after using ADASYN less category synthesis method, the evaluation indexes of BiGRU network are higher than BiLSTM network under the same parameters, so BiGRU network should be assigned a higher overall weight. The original samples are trained with the same network structure and parameters. The test results are shown in Table 1. It can be seen that after ADASYN synthesizes a small number of samples, the classification indexes of the two neural networks are significantly improved, the first-class rating indexes of BiGRU network with good performance are increased by nearly 15%, and all evaluation indexes of BiGRU network are higher than those of BiLSTM network. It is further concluded that the performance of BiGRU is better than BiLSTM, and a higher overall weight can be assigned to BiGRU network.

4.2   Weight calculation of BiGRU and BiLSTM

In order to more comprehensively obtain the performance of neural network in each category classification, a few category samples synthesized by ADASYN and all real samples are used. A total of 6327 samples are input into the trained ADASYN + BiGRU and ADASYN + BiLSTM neural networks. The category weight calculation results of the two neural networks in the primary classification are shown in Table 2.

It can be seen from Table 2 that although BiGRU has higher overall evaluation index and higher overall weight than BiLSTM, the performance of the two neural networks are different in each category. BiLSTM has a larger category weight in the categories of security inspector, public works equipment and unknown reason, indicating that BiLSTM network has decision-making power in these three categories. Due to the large number of secondary classification categories of signal turnout equipment faults, considering the length, this paper only lists the weight calculation results of primary classification categories.

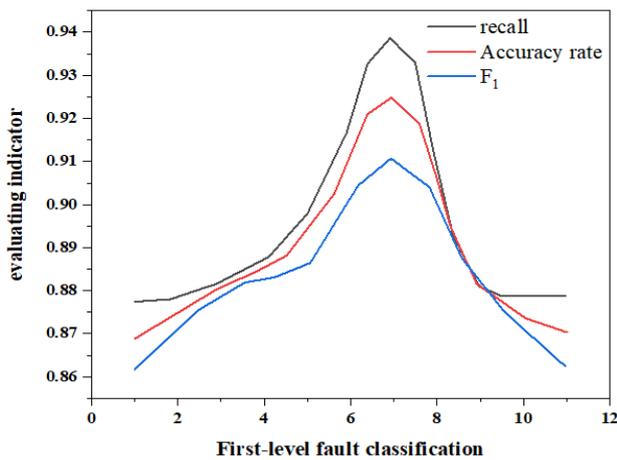## 4.3   Deep learning integration model and classification

The various weights of the neural network are obtained through the above tests, and BiGRU should have higher overall weight than BiLSTM. Different overall weights are given to BiGRU and BiLSTM. The two deep learning neural networks are integrated through combined weighting, and the common classification prediction results are obtained through recalculation of the outputs of the two networks.

Under different overall weight distribution, see Figure 6 for the evaluation indexes of level 1 fault classification and level 2 fault classification of the deep learning integration model (where G represents BiGRU and L represents BiLSTM). It can be seen from Figure 6 that when the overall weight of BiGRU is 0.54 and the overall weight of BiLSTM is 0.46, the evaluation index of the deep learning integration model is the highest. The

final classification results of the deep learning integration
model are shown in Table 3 and Figure 7.

Classification	Classification method	Number of classification errors / total number of categories	Recall rate	Category weight
Switch machine	ADASYN＋BiGRU	266/2053	0.1295	1.9048
	ADASYN＋BiLSTM	288/2053	0.1403	1.8129
External locking and installation device	ADASYN＋BiGRU	163/1251	0.1303	1.8983
	ADASYN＋BiLSTM	192/1251	0.1534	1.7076
Paste checker	ADASYN＋BiGRU	81/567	0.1428	1.7918
	ADASYN＋BiLSTM	70/567	0.1235	1.9601
Turnout control circuit equipment	ADASYN＋BiGRU	167/1280	0.1305	1.8968
	ADASYN＋BiLSTM	189/1280	0.1477	1.7531
Permanent way equipment	ADASYN＋BiGRU	62/440	0.1409	1.8077
	ADASYN＋BiLSTM	55/440	0.1250	1.9459
Supporting equipment	ADASYN＋BiGRU	86/614	0.1401	1.8147
	ADASYN＋BiLSTM	80/614	0.1303	1.8984
Unknown reason	ADASYN＋BiGRU	21/124	0.1694	1.5902
	ADASYN＋BiLSTM	14/124	0.1129	2.0614

Table 2: Calculation results of class I classification weight of signal turnout equipment fault

(a): First-level fault classification
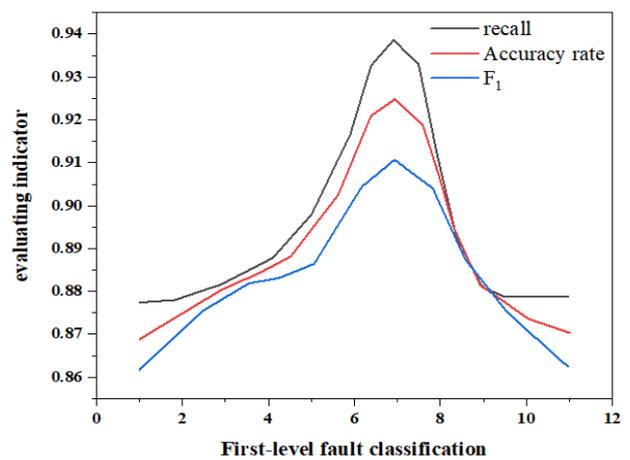
(b): BiLSTM secondary classification training process

Figure 6: Evaluation index values of deep learning integration model under different overall weight distribution

Method	Level	Accuracy rate	Recall rate	F1 value
Deep learning integration model	Primary fault classification	0.9106	0.9389	0.9245
	Secondary fault classification	0.8564	0.8612	0.8588

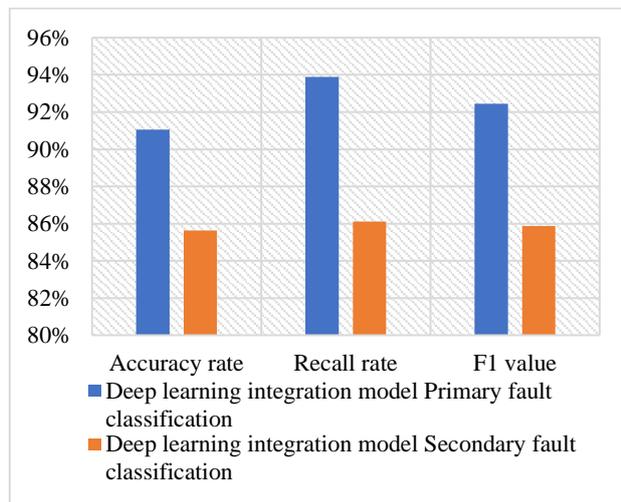Table 3: Classification test results of deep learning integration model

Figure 7: Graphical representation of classification test results of deep learning integration model

It can be seen from Table 3 and Figure 7 that compared with ADASYN + BiGRU neural network, the comprehensive evaluation index of primary fault classification is improved by about 5%, and the comprehensive evaluation index of secondary fault classification is improved by about 9%. Compared with ADASYN + BiLSTM neural network, the comprehensive evaluation index of primary fault classification is improved by about 6%, and the comprehensive evaluation index of secondary fault classification is improved by about 10%.

5   Conclusions

This paper studies the fault diagnosis model of signal turnout fault text data, uses ADASYN data synthesis method to synthesize few category samples. This article also uses TF-IDF to extract features and transform vectors, and puts forward a deep learning integration method based on combination weight. The sample synthesis algorithm can appropriately synthesize a few categories of samples according to the distribution of the overall samples. There are several articles which uses SVM-SMOTE method to automatically synthesize the few category samples of signal equipment fault, and solve the problem of signal equipment fault sample imbalance. Through experimental analysis, it is proved that deep learning integration is a method that can effectively improve the classification performance of turnout fault diagnosis model. At the same time, this method can also provide a new idea for railway text classification and fault diagnosis. This article utilizes the deep learning algorithm of artificial intelligence for investigating the interlocking faults in the railway transportation. This paper uses ADASYN data synthesis method to synthesize few category samples, uses TF-IDF to extract features and transform vectors, and proposes a deep learning integration method based on combined weight. The outcomes obtained for the proposed method reveals that BiGRU has better overall classification performance when evaluated on the index of primary and secondary fault classification accuracy.

References

[1] Kong, J. (2020, December). Application and research of artificial intelligence in digital library. In *International conference on Big Data Analytics for Cyber-Physical-Systems* (pp. 318-325). Springer, Singapore.
https://doi.org/10.1007/978-981-33-4572-0 47

[2] Paek, S., & Kim, N. (2021). Analysis of worldwide research trends on the impact of artificial intelligence in education. *Sustainability*, *13*(14), 7941.
https://doi.org/10.3390/su13147941

[3] Dobias, R., & Kubatova, H. (2004, August). FPGA based design of the railway's interlocking equipments. In *Euromicro Symposium on Digital System Design, 2004. DSD 2004.* (pp. 467-473). IEEE.
10.1109/DSD.2004.1333312

[4] Skiribou, C., Elbahhar, F., & Elassali, R. (2021). DMRS-based channel estimation for railway communications in tunnel environments. *Vehicular Communications*, *29*, 100340.
https://doi.org/10.1016/j.vehcom.2021.100340

[5] Kiedrowski, P., & Saganowski, Ł. (2021). Method of Assessing the Efficiency of Electrical Power Circuit Separation with the Power Line Communication for Railway Signs Monitoring. *Transport and Telecommunication*, *22*(4), 407-416.
10.2478/ttj-2021-0031

[6] Yang, J., Bai, X., Zhang, Z., Yang, M., Pan, P., Liu, T., & Tao, T. (2021, May). Research on the application of BDS/GIS/RS technology in the high speed railway infrastructure maintenance. In *IOP Conference Series: Earth and Environmental Science* (Vol. 783, No. 1, p. 012168). IOP Publishing.
10.1088/1755-1315/783/1/012168

[7] Lin, J., Hu, X., Dang, J., & Wu, Z. (2019). Traffic model of machine-type communication for railway signal equipment based on MMPP. *IET Microwaves, Antennas & Propagation*, *13*(8), 1072-1079.
https://doi.org/10.1049/iet-map.2018.6004

[8] Wang, X., Guo, J., Jiang, L., Fu, J., & Li, B. (2016, August). Intelligent fault diagnosis and prediction technologies for condition based maintenance of track circuit. In *2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT)* (pp. 276-283). IEEE.
10.1109/ICIRT.2016.7588745

[9] Sahal, R., Breslin, J. G., & Ali, M. I. (2020). Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. *Journal of manufacturing systems*, *54*, 138-151.
https://doi.org/10.1016/j.jmsy.2019.11.004

[10] Cao, Y., Li, P., & Zhang, Y. (2018). Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing. *Future Generation Computer Systems*, *88*, 279-283.
https://doi.org/10.1016/j.future.2018.05.038

[11] Dong, C. Z., Ye, X. W., & Jin, T. (2018). Identification of structural dynamic characteristics based on machine vision technology. *Measurement*, *126*, 405-416.
https://doi.org/10.1016/j.measurement.2017.09.043

[12] Jia, Z., & Sharma, A. (2021). Review on engine vibration fault analysis based on data mining. *Journal of Vibroengineering*, *23*(6), 1433-1445.
https://doi.org/10.21595/jve.2021.21928

[13] Yin, M., Li, K., & Cheng, X. (2020). A review on artificial intelligence in high-speed rail. *Transportation Safety and Environment*, *2*(4), 247-259.
https://doi.org/10.1093/tse/tdaa022

[14] Ren, X., Li, C., Ma, X., Chen, F., Wang, H., Sharma, A., & Masud, M. (2021). Design of multi-information fusion based intelligent electrical fire detection system for green buildings. *Sustainability*, *13*(6), 3405.
https://doi.org/10.3390/su13063405

[15] Sharma, D., Kaur, R., Sandhir, M., & Sharma, H. (2021). Finite element method for stress and strain analysis of FGM hollow cylinder under effect of temperature profiles and inhomogeneity parameter. *Nonlinear Engineering*, *10*(1), 477-487.
https://doi.org/10.1515/nleng-2021-0039

[16] Afandizadeh, S., & Rad, H. B. (2021). Developing a model to determine the number of vehicles lane changing on freeways by Brownian motion method. Nonlinear Engineering, 10(1), 450-460.
https://doi.org/10.1515/nleng-2021-0036

[17] Shabaz, M., Sharma, A., Al Ajrawi, S., & Estrela, V. V. (2022). Multimedia-based emerging technologies and data analytics for Neuroscience as a Service (NaaS). Neuroscience Informatics, 2(3), 100067.
https://doi.org/10.1016/j.neuri.2022.100067

[18] Meher, M., & Rostamy, D. (2021). Hybrid of differential quadrature and sub-gradients methods for solving the system of Eikonal equations. Nonlinear Engineering, 10(1), 436-449.
https://doi.org/10.1515/nleng-2021-0035

[19] Mi, Z., Wang, T., Sun, Z., & Kumar, R. (2021). Vibration signal diagnosis and analysis of rotating machine by utilizing cloud computing. Nonlinear Engineering, 10(1), 404-413.
https://doi.org/10.1515/nleng-2021-0032

[20] Wang, H., Sharma, A., & Shabaz, M. (2022). Research on digital media animation control technology based on recurrent neural network using speech technology. International Journal of System Assurance Engineering and Management, 13(1), 564-575.
https://doi.org/10.1007/s13198-021-01540-x

[21] Yousaf, B., Qaisrani, M. A., Khan, M. I., Sahar, M. S. U., & Tahir, W. (2021). Numerical and

experimental analysis of the cavitation and study of flow characteristics in ball valve. Nonlinear Engineering, 10(1), 535-545. *https://doi.org/10.1515/nleng-2021-0044*

[22] Singh, P. K., & Sharma, A. (2022). An intelligent WSN-UAV-based IoT framework for precision agriculture application. *Computers and Electrical Engineering*, *100*, 107912. *https://doi.org/10.1016/j.compeleceng.2022.107912*

[23] Zeng, H., Dhiman, G., Sharma, A., Sharma, A., & Tselykh, A. (2021). An IoT and Blockchain-based approach for the smart water management system in agriculture. *Expert Systems*, e12892. *https://doi.org/10.1111/exsy.12892*

[24] Sharma, A., & Singh, P. K. (2021). UAV-based framework for effective data analysis of forest fire detection using 5G networks: An effective approach towards smart cities solutions. *International Journal of Communication Systems*, e4826. *https://doi.org/10.1002/dac.4826*

[25] Sharma, A., Singh, P. K., & Kumar, Y. (2020). An integrated fire detection system using IoT and image processing technique for smart cities. *Sustainable Cities and Society*, *61*, 102332. *https://doi.org/10.1016/j.scs.2020.102332*

[26] Zang, Y., Shangguan, W., Cai, B., Wang, H., & Pecht, M. G. (2019). Methods for fault diagnosis of high-speed railways: A review. Proceedings of the institution of mechanical engineers, part O: journal of risk and reliability, 233(5), 908-922. *https://doi.org/10.1177/1748006X18823932*

[27] Ting, L., Khan, M., Sharma, A., & Ansari, M. D. (2022). A secure framework for IoT-based smart climate agriculture system: Toward blockchain and edge computing. Journal of Intelligent Systems, 31(1), 221-236. *https://doi.org/10.1515/jisys-2022-0012*

[28] Minea, M., Dumitrescu, C. M., & Dima, M. (2021). Robotic Railway Multi-Sensing and Profiling Unit Based on Artificial Intelligence and Data Fusion. Sensors, 21(20), 6876. *https://doi.org/10.3390/s21206876*

[29] Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big data application in biomedical research and health care: a literature review. *Biomedical informatics insights*, *8*, BII-S31559. *https://doi.org/10.4137/BII.S31559*

[30] Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V., & Rong, X. (2015). Data mining for the internet of things: literature review and challenges. *International Journal of Distributed Sensor Networks*, *11*(8), 431047. *https://doi.org/10.1155/2015/431047*