Proceedings of Machine Learning Research vol 245:1–17, 2024 2024 International Conference on Machine Learning and Intelligent Computing

Fusion Analysis and Digital Realization of Inspection and Repair of High-Speed Railway Engineering Equipment

Xiaoai Zhou 18706406901@163.COM
Institute of Computing Technology, China Academy of Railway Sciences, Beijing, China

Editors: Nianyin Zeng and Ram Bilas Pachori

Abstract

High-speed railway engineering equipment is the basis of railway transportation, and transportation safety is closely related to the equipment operating state. In order to help inspectors and managers of the engineering departments to conduct inspection and repair more efficiently, solve problems such as the lack of intuitiveness of inspection status, unreasonable setting of inspection cycle days, and frequent occurrence of equipment accidents caused by seasonal weather changes, this paper uses deep learning to establish an automatic disease recognition model based on Convolutional Neural Network. Through the data collected from several high-speed railway workshops for verification, it is concluded that the model can realize the automatic recognition of disease types, the training accuracy reaches 97%, and the verification accuracy reaches 76%. Meanwhile, based on big data technology, this paper combines Convolutional Neural Network and Long Short-Term Neural Network, establishes the equipment status judgment model, and builds the inspection cycle algorithm based on equipment status. Through the data collected from multiple engineering departments, the equipment state judgment model can capture the key information from the inspection records and can thus accurately judge the equipment operating status. The accuracy of the predicted inspection times reaches 85.7%. Finally, through the digital implementation case, it is fully proved that the design and fusion of these two applications can provide sufficient technical support for the inspection and repair of high-speed railway engineering equipment and can thus provide a comprehensive and reliable data basis for real-time and efficient inspection and repair decisions.

Keywords: High-Speed Railway Engineering Equipment, Inspection and Repair, Convolutional Neural Network, Automatic Disease Recognition Model, Early Warning State, Cycle Days.

1. Introduction

The engineering equipment of high-speed railway is the important foundation of railway transportation, and the safety of railway transportation is closely related to the using state of the engineering equipment. With the rapid growth of high-speed railway operating mileage, the transportation department has formulated higher standards for the ledger management, disease inspection, and disease maintenance of engineering equipment (Yang et al., 2018). Although the engineering major of high-speed railway has accumulated a great amount of equipment ledger data and maintenance data, and has realized the digital management of engineering equipment maintenance through information technology to a certain extent, the intelligence level of inspection management of railway equipment is still low compared with that of foreign countries.

In terms of equipment management, the periodic inspection status of domestic engineering equipment is not intuitive; the dynamic correlation between the inspection cycle and the operating state of engineering equipment has not been realized; and the digital management of spring-and-autumn special inspection has not been realized either. In terms of periodic inspection, more and more high-speed railway engineering departments have realized the digital management of periodic inspection (Shi, 2019), which effectively realized the automatic generation of inspection plans according to the number of days in the inspection cycle, the online distribution of periodic inspection tasks, and the online upload of inspection records by means of information technology, and thus provided technical support for on-site staffs at all levels to track the progress and results of periodic inspection.

However, at present, many inspection and repair management systems lack the early warning status of periodic inspection items of engineering equipment. Taking the engineering safety production management system as an example, users cannot intuitively find which equipment or mileage range have overdue inspection items. They can only search the latest inspection records in the system and manually com-pare the latest inspection date with the current date to determine whether the inspection items are overdue or approaching overdue. Moreover, due to the weak visibility of early warning status, users cannot judge which type of engineering equipment is not timely nor standardized in inspection by comparing the number of overdue early warnings of various equipment types. The lack of intuitiveness and statistical analysis of early warning states reduces the decision-making efficiency and accuracy of the leaders and workers in on-site engineering stations. In addition, the period days of each inspection item are set manually, and the number of inspection period days cannot be dynamically adjusted according to the operating status of the engineering equipment. The periodic inspection management system is not fully connected with the equipment ledger, and it cannot realize the bidirectional synchronization nor the fusion analysis of the equipment status information and the periodic inspection plan setting, and it is difficult to adjust the inspection and repair progress and strategy in time when there is a change in the equipment status. Meanwhile, due to seasonal weather changes and other factors, the spring and autumn each year are the most prone to equipment accidents in the high-speed railway engineering departments. However, inspectors can only record detailed inspection and disease information by recording electronic forms or by hand.

In recent years, machine learning and big data have been widely applied in the field of high-speed railway in foreign countries like the United States or Japan. Struckon Company in the United States has established a prediction model of equipment rail failure risk by using data clustering, and Japan has realized automatic identification of equipment status and fault by using artificial intelligence (Zheng et al., 2018). Intuitive early warning information can provide key technical support for field staff to efficiently compare and analyze the periodic inspection progress and standardization of various engineering equipment, and can provide an intuitive data basis for the adjustment of periodic inspection plans and equipment management decisions. The bidirectional synchronization between the number of inspection cycle days and the operating state of the equipment can further improve the accuracy of the periodic inspection and repair decision. At the same time, the highly visualization of disease records and the electronic closed-loop audit process can provide essential technical support to help the inspectors and managers of the engineering departments to improve the efficiency of inspection and repair and to ensure the timeliness and accuracy of decision-making.

Therefore, the digital technology scheme is researched and designed to realize the visual display and statistical analysis of the early warning information of each inspection item and the automatic correlation between inspection cycle days and equipment status. Meanwhile, it is necessary to design the information technology scheme to realize the comprehensive management of spring-and-autumn special inspection records.

2. Automatic Disease Recognition Model for Spring-and-Autumn Special Inspection and Repair

Spring-and-autumn special inspection and repair content varies according to the state of the line; its inspection items include line static inspection, ballastless track structure inspection, rail disease inspection, switch rail wear inspection, curve rail wear inspection. The diseases of engineering equipment mainly include nine categories: wear, falling, crack, seam, abrasion, rust, fish scale, cross section, and loose. In this paper, an automatic disease recognition model based on Convolutional Neural Network has been proposed.

Convolutional Neural Networks (CNN) are widely used in the field of image recognition and classification, aiming to learn text feature representations from two-dimensional images (Pan et al., 2024). This kind of neural network can be divided into five layers: input layer, convolutional layer, pooling layer, fully connected layer, and output layer.

In this study, the disease images uploaded by inspectors during special inspection in spring and autumn can be directly input into the convolutional neural network as input data. The design of the disease recognition model is based on CNN (see Fig. 1), and the digital units are all pixels. The model consists of five convolution layers, five pooling layers, one flatten layer, and three fully connected layers.

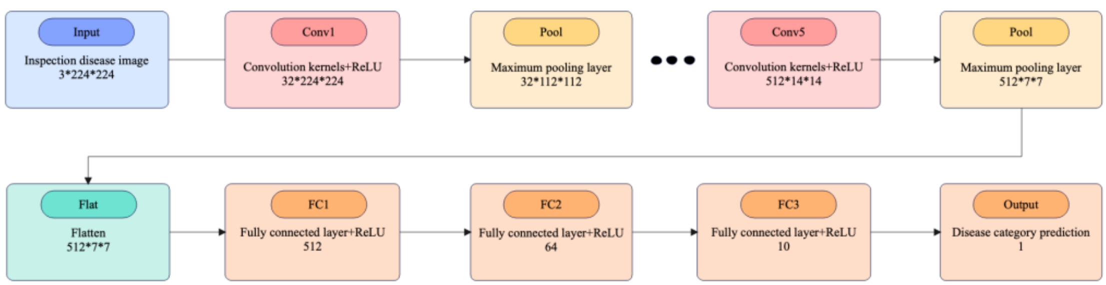

Figure 1: Neural network architecture of the automatic disease recognition model

Each convolution layer contains 32 convolution kernels of size [3, 3]. Due to the different weights of convolution kernels in each dimension and the local positions targeted, each convolution kernel can capture different features (K, Subhashini et al., 2023). Thus, each convolution layer can obtain 32 characteristics of the disease. The images are first convolved, and then the nonlinear features are extracted, and the feature matrix is out-put by activation function. This model uses ReLU as the activation function, expressed by formula (1).

$$\mathrm{ReLU}(x) = \max(0, x) \tag{1}$$

The maximum pooling layer of size [2, 2] is used to reduce the dimension of the feature matrix and screen out the most significant features. The calculation of the pixel size of the output image of the maximum pooling layer follows formula (2), where i represents the size of the input image, k represents the size of the pooling core, and s represents the size of the pooling core sliding window. The pooling layer used in this model reduces the height and the width of the feature matrix by half.

$$\left\lfloor \frac{i-k}{s} \right\rfloor + 1 \tag{2}$$

The flatten layer transforms the multi-dimensional feature matrix processed by the pooling layer into a one-dimensional vector and transitions to the fully connected layer. The fully connected layer integrates local features and uses Softmax activation formula (3) to map the image features extracted through convolution and pooling into the disease category matrix mapped as n*1.

$$\mathrm{Softmax}\,(x)_i = \frac{\mathrm{e}^{x_i}}{\sum\limits_{k=1}^{\mathrm{K}} \mathrm{e}^{x_k}} \tag{3}$$

3. Periodic Inspection Equipment State Judgment Model Considering Dynamic Synchronization

The operating state of engineering equipment is mainly divided into in use, overhaul, update, and scrap. To determine the running status of a device based on the detailed inspection record associated with the device is a necessary basis for dynamic syn-chronization of the inspection period and equipment status. In this paper, a periodic inspection equipment state recognition model considering dynamic synchronization has been proposed.

3.1. Equipment State Judgment Model Based on Big Data Technology

Although the detailed records of periodic inspection and repair uploaded by inspec-tors, including construction contents, equipment changes, operation plans, and prob-lem descriptions, can provide an important basis for the judgment of equipment sta-tus, the information of inspection and repair records is complex and lengthy, and it is impossible to effectively extract keywords from them to judge the operating status of related equipment. Therefore, a hybrid model based on Convolutional Neural Net-work (CNN) and Long Short-Term Memory artificial neural network (LSTM) has been introduced in this study.

CNN can capture the text features of sentences through convolution kernels of different sizes, and then carry out text classification. The two-dimensional vector expression of the text serves as the input layer of the CNN model. CNN will extract text features of different levels in its convolution layer, filter information in the pooling layer, associate different text features in the fully connected layer, map them into a one-dimensional text category matrix (Shin et al., 2018; Ren, 2023), and finally realize text classification in the output layer. Different from CNN, LSTM (Long Short-Term Memory) is a special recurrent neural network (RNN), which is widely used in the interpretation and processing of long text. LSTM is composed of long-term memory and short-term memory, including three gates: forgetting gate, input gate, and output gate. The for-getting gate determines which information should be forgotten, the input gate deter-mines which new information should be incorporated into the cellular state, and the output gate determines which should constitute short-term memory in the updated long-term memory (Zhang et al., 2023). The mathematical expression of the gate structure is shown in formulas (4) to (8).

$$\mathrm{f}_t = \sigma\left(\mathrm{W_f} \cdot \left[\mathrm{h}_{t-1}, \mathrm{x}_t\right] + \mathrm{b_f}\right) \tag{4}$$

$$\mathrm{i}_t = \sigma\left(\mathrm{W}_{\mathrm{i}} \cdot \left[\mathrm{h}_{t-1}, \mathrm{x}_t\right] + \mathrm{b}_{\mathrm{i}}\right) \tag{5}$$

$$\mathrm{C}_t = \tanh\left(\mathrm{W}_{\mathrm{C}} \cdot \left[\mathrm{h}_{t-1}, \mathrm{x}_t\right] + \mathrm{b}_{\mathrm{C}}\right) \tag{6}$$

$$\mathrm{o}_t = \sigma\left(\mathrm{W}_{\mathrm{o}} \cdot \left[\mathrm{h}_{t-1}, \mathrm{x}_t\right] + \mathrm{b}_{\mathrm{o}}\right) \tag{7}$$

$$\mathrm{h}_t = \mathrm{o}_t \cdot \tanh\left(\mathrm{C}_t\right) \tag{8}$$

f_t, i_t, C_t, o_t, and h_t represent the cell state of forgetting gate, input gate, intermediate quantity, and output gate at time t, respectively. σ represents the activation function Sigmoid; W represents the learnable weight of each gate; b represents the offset entry of each input gate; h_{t-1} and h_t represent the hidden layer state at t-1 and t moments, respectively; x_t represents the current input. Through this ”gate” structured memory mechanism, LSTM is able to maintain a long-term memory state and efficiently capture contextual dependencies.

In this study, a periodic inspection record corresponds to a text as input data. The architecture is based on CNN-LSTM (see Fig. 2). The CNN part of the model contains fifteen convolution kernels of different sizes, a maximum pooling layer, and a fully connected layer, which are used to extract features of phrases of different lengths. The number of LSTM layers is set to 1 and the hidden layer dimension is set to 128. The features obtained by CNN and LSTM are fused and spliced through a fully connected layer, and the equipment status is output in the output layer as the final classification result.

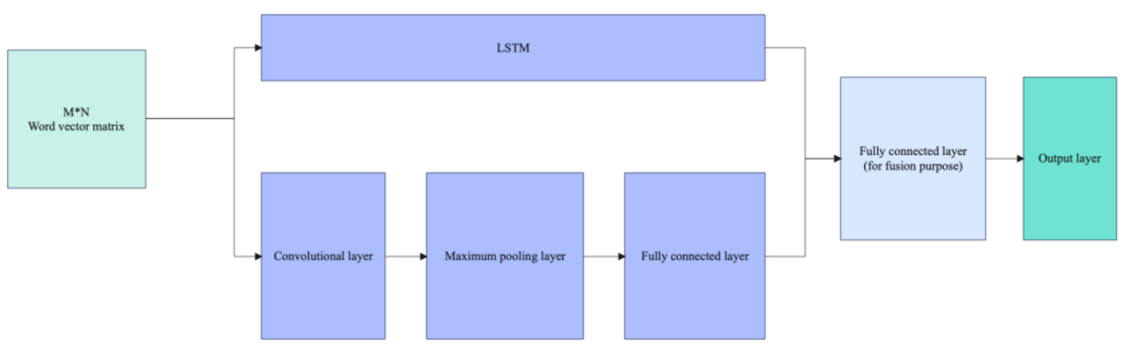

Figure 2: Neural network architecture of the periodic inspection equipment state judgment model

The preprocessing of text classification data aims to transform natural language into word vectors that can be recognized by computers. Word2Vec is used in this study. Word2Vec is a lightweight neural network, which adopts a distributed coding method and can effectively avoid the data disaster brought by the traditional One-Hot coding method (Xie, 2024). Word2Vec contains two models, CBOW and Skip-Gram (see Fig. 3 and Fig. 4). The CBOW model predicts the target word based on the context, while the Skip-Gram model predicts the context based on the target word. These two models are combined to realize the conversion of discrete words to dense vectors (Peng et al., 2021).

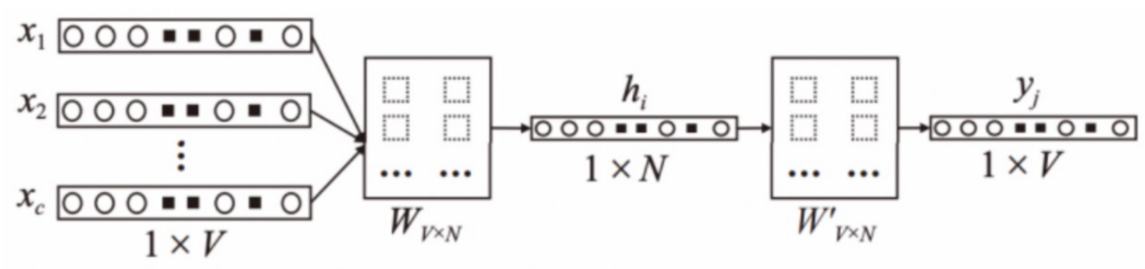

Figure 3: CBOW model architecture (Xie, 2024)

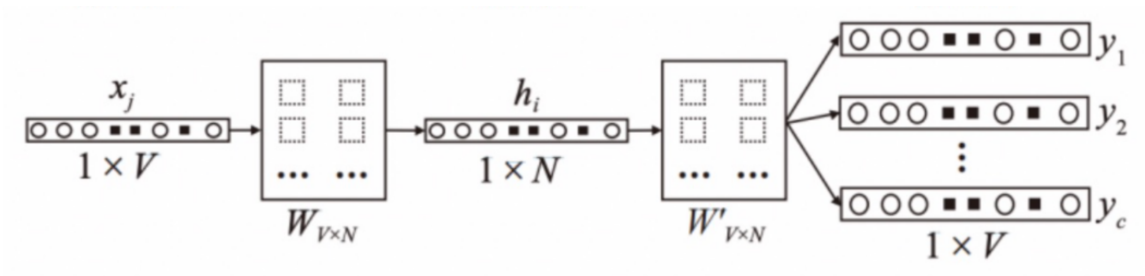

Figure 4: Skip-Gram model architecture (Xie, 2024)

The word vectors obtained from the text data processed by Word2Vec constitute the data set for this model training. Cross entropy is used as the loss function and RAdam (Rectified Adam) is used as the optimizer. Compared with the traditional Adam optimizer, RAdam is able to modify the variance, and combine with the Momentum algorithm of stochastic gradient descent (SGD Momentum) to realize the automatic adjustment of hyperparameters in the training stage, so that the model can achieve higher training efficiency and accuracy (Huang et al., 2023). The following Table 1 gives the initial setting of hyperparameters.

Table 1: Equipment state judgment model hyperparameter initialization

Hyperparameters	Value
Batch size	32
Learning rate	1.00E-5
Number of epochs	30

In this paper, the confusion matrix is used to evaluate the accuracy of the equipment state recognition model. The confusion matrix is divided into four quadrants: TP (true positive), FN (false negative), FP (false positive) and TN (true negative), including accuracy rate, recall rate, and F1 score. The calculation method is shown in formulas (9) to (11).

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{9}$$

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{10}$$

$$F1 = \frac{2 \times (P \times TPR)}{P + TPR} \tag{11}$$

The accuracy rate represents the proportion of positive samples predicted by the model. Recall rate represents the proportion of samples that can be correctly predict-ed as positive by the model. The F1 score is the harmonic average of accuracy and recall (Huang et al., 2024).

The macro average model is used to evaluate the multi-class classification model of the equipment state judgment model. The calculation formula of the quantitative index is shown in formulas (12) to (14).

$$M_P = \frac{1}{N}\sum_{i=1}^{N} P_i \tag{12}$$

$$M_{TPR} = \frac{1}{N}\sum_{i=1}^{N} TPR_i \tag{13}$$

$$M_{F1} = \frac{1}{N}\sum_{i=1}^{N} F1_i \tag{14}$$

In these formulas, M_P denotes the macro average accuracy, M_{TPR} denotes the macro average recall rate, and M_{F1} denotes the macro average F1 score.

3.2. Construction of Inspection Cycle Algorithm Based on Equipment State

Now, the automatic judgment of equipment status has been realized through big data technology, and in order to realize the bidirectional synchronization between the inspection cycle and the equipment status, the only remaining step is the mapping of the two. First, this study proposes the mapping relationship shown in Table 2. n indicates the number of inspection cycles of the equipment in use. When the equipment is in the overhaul state, the inspection cycle must be shortened to achieve comprehensive inspection and maintenance of the equipment performance. When the up-dated state is reached, the frequency of diseases decreases due to technological innovation and the elimination of damaged equipment, and thus the frequency of inspection will accordingly decrease in the short term.

Table 2: Mapping table between the inspection period and equipment status

Equipment status	Inspection period days
In use	n
Overhaul	1.5n
Update	0.75n
Scrap	∞

In order to realize the mapping from equipment state to inspection period, an algorithm based on a hash table has been proposed. Hash tables are strong in fast lookup of elements. For this model, this data structure can directly access the corresponding value ”inspection period days” according to the keyword "equipment status", the time complexity is only O(1), and the system memory is greatly saved.

In this model, the keyword "equipment state" field is denoted as k, and the hash table length is m. The divisor method is used to construct the hash function (Huang et al., 2022). The dividend is the keyword k, and the divisor is the prime p not greater than the table length m. The hash function obtained is shown in formula (15). It should be noted that the keywords of the hash table are generally in integer form. Since the key of this model is a string, it needs to be converted into a numeric value through ASCII code and then the hash function can be constructed.

$$\mathrm{H}(k) = k\mathrm{MODp} \tag{15}$$

Through the combination of the automatic recognition model of equipment status and the mapping model of equipment status and inspection period, the dynamic correlation between inspection period and the operating state of engineering equipment can be realized. The system can automatically adjust the maintenance period when the equipment status changes, and thus provides a reliable basis for field work-ers to timely adjust the periodic inspection and repair plan and strategy.

3.3. Recommended Inspection Date Estimation Model Based on Mathematical Function Model

In the actual inspection process, in order to avoid railway safety hazards caused by not timely maintenance, the concept of "more inspection rather than late inspection" is always adhered to for engineering projects with short inspection cycles. For exam-ple, the period of switch fault inspection is 30 days, but in fact, the engineering departments may inspect once a week. In this case, the periodic inspection early warn-ing mentioned above can only provide necessary but insufficient support for on-site workers to make inspection and repair decisions. In view of this limitation, a recommended inspection date prediction mechanism based on GM time series model has been proposed to further optimize the inspection early warning system.

GM (1,1) model is a dynamic prediction model, which arranges the predicted objects in chronological order and quantitatively predicts their future development rules by analyzing the historical development rules of the predicted objects. It is suitable for the research of problems with less original data and irregular sample sequence (Wang et al., 2024; Islam et al., 2022). In this study, GM (1,1) model predicts the number of inspections in the next quarter by analyzing the trend of the number of inspections in each quarter of a specific inspection item in the past three years, which lays the foundation for the estimation of the recommended inspection date. The solution of the model varies according to the specific equipment type and inspection items, and the switch fault inspection is taken as an example here.

The number of switch fault detection intervals in each quarter from January 2021 to March 2024 is selected as the sample to construct the model. The mathematical principle is applied as follows:

The original number sequence is listed in formula (16), which is arranged according to the time series.

$$\mathrm{X}^{(0)}(t) = \left\{\mathrm{X}^{(0)}(1), \quad \mathrm{X}^{(0)}(2), \mathrm{X}^{(0)}(\mathrm{n})\right\} \tag{16}$$

The original sequence is summed once to obtain the formula (17).

$$\mathrm{X}^{(1)}(t) = \left\{\mathrm{X}^{(1)}(1), \quad \mathrm{X}^{(1)}(2), \quad \mathrm{X}^{(1)}(\mathrm{n})\right\} \tag{17}$$

A sequence of adjacent means is generated through $\mathrm{Z}^{(1)}(t) = \frac{1}{2}\left[\mathrm{X}^{(1)}(t) + \mathrm{X}^{(1)}(t-1)\right]$, as shown in formula (18).

$$Z^{(1)}(t) = \left\{Z^{(1)}(1), \quad Z^{(1)}(2), \quad Z^{(1)}(\mathrm{n})\right\} \tag{18}$$

The first order ordinary differential equation $\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t} + \mathrm{a}x^{(1)} = \mathrm{b}$ is construct-ed by using the least square method. The parameters a and b are constructed as shown in formula (19). a represents the development coefficient and b represents the grey action. Y and B are represented in formulas (20) and (21), respectively.

$$(\mathrm{a}, \mathrm{b})^{\mathrm{T}} = (\mathrm{BB})^{-1}\mathrm{B}^{\mathrm{T}}\mathrm{Y} \tag{19}$$

$$\mathrm{Y} = \begin{bmatrix} \mathrm{X}^{(0)}(2) \\ \mathrm{X}^{(0)}(3) \\ \mathrm{X}^{(0)}(\mathrm{n}) \end{bmatrix} \tag{20}$$

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ \dots & \cdots \\ -Z^{(1)}(n) & 1 \end{bmatrix} \tag{21}$$

The prediction formula of GM (1,1) model is obtained, as shown in formula (22).

$$\mathrm{X}^{(1)}(t+1) = \left[\mathrm{X}^{(0)}(1) - \frac{\mathrm{b}}{\mathrm{a}}\right] \mathrm{e}^{-\mathrm{a}(t-1)} + \frac{\mathrm{b}}{\mathrm{a}} \tag{22}$$

The estimated recommended inspection date is completely based on the above quarterly inspection frequency estimation mechanism, which is briefly summarized as the following: use the current date plus the estimated inspection cycle days of the current quarter to obtain the recommended date of the next inspection, as shown in formula (23).

$$\mathrm{T}(i+1) = \mathrm{T}(i) + \frac{\mathrm{D}(i+1)}{\mathrm{N}(i+1)} \tag{23}$$

In formula (23), $\mathrm{T}(i+1)$ represents the next inspection date, $\mathrm{T}(i)$ represents the last inspection date, $\mathrm{D}(i+1)$ represents the number of days in the current quarter, and $\mathrm{N}(i+1)$" " represents the estimated number of inspections in the current quarter.

For example, based on the GM (1,1) inspection frequency prediction model, the inspection frequency of switch fault detection items in the second quarter of 2024 is 23, and the corresponding recommended inspection period is therefore 3 days. There-fore, although the specified inspection period of switch fault is 30 days, the actual inspection period affected by various factors is three days instead. After the recommended inspection period, the system will automatically display the suggested inspection pop-up window to prompt on-site workers to conduct inspection. In this example, it is assumed that the last inspection date of switch detection is April 1. Although the early warning status of the project stays "normal" before May 1, based on the suggested inspection date estimation mechanism, the system will provide inspection reminders on April 4.

4. Model Solution and Numerical Example

4.1. Model Solution and Numerical Example of the Automatic Disease Recognition Model

The data of this model was collected from several workshops of a high-speed rail-way. A total of 12,000 images of each type of disease and 12,000 disease-free images were collected, and the image size was 224*224. The images of each disease category were divided into a training set, a verification set, and a test set according to the ratio of 8:1:1. The same equipment under each disease category has obvious differences in geometric shape and color, so the CNN model can classify the diseases ac-cording to the geometric shape and the color of the equipment.

CNN can explore the mapping relationship between the learning sample and the target sample by the size of the parameter weights. In this training, the forward and backward iteration method is used to adjust the weights of learning parameters. The forward iteration calculates the result and makes the prediction based on the input and the set parameters and functions. In backward iteration, parameters are learned and optimized according to the gradient value given by the loss function (Shi et al., 2022). In this model, cross entropy, the most widely used in the multi-class classification problem, is adopted as the loss function, and the expression is formula (24), where N represents the number of training samples, which is 96,000 in this training. K represents the number of categories, which in this training is 10.

$$\mathrm{CE} = -\frac{1}{\mathrm{N}} \sum_{n=1}^{\mathrm{N}} \sum_{k=1}^{\mathrm{K}} t_{n,k} \log\left(y_n, k\right) \tag{24}$$

This training uses the Adam optimizer, under which each weight has a different rate, to adjust the model's adaptive learning rate through faster convergence. Hyperparameters are another important feature of model training, which determines the weight and rate of learning. In this training, hyperparameters include the number of samples processed in each batch, the learning rate, and the number of iterations. The initial values are set as shown in Table 3.

Table 3: Disease recognition model hyperparameter initialization

Hyperparameters	Value
Batch size	25
Learning rate	1.00E-3
Number of epochs	30

The adjustment of hyperparameters directly affects the efficiency and accuracy of the disease recognition model, and the adjustment process is also an essential part in the training process. In order to minimize the degree of overfitting and maximize the accuracy of verification, the following hyperparameter adjustment mechanism is used in this training, and the accuracy and loss curve of training and verification are adopted as the adjustment basis: If the verification accuracy is low, batch size will be increased to improve the generalization degree of the model; if the convergence rate of the training curve is slow, the learning rate will be increased; and if the training curve lacks stability and the noise is obvious, the learning rate will be reduced. Ac-cording to the accuracy trend of the last several iterations in the verification curve, if the learning potential of the model is shown, the number of epochs will be increased. Otherwise, if the overfitting is shown, the number of iterations will be reduced.

After continuous adjustment and optimization of hyperparameters, the batch size of the final disease recognition model was set to 32, the learning rate to 0.001, and the number of epochs to 27. The final training accuracy is as high as 97%, and the verification accuracy reaches 76% (see Fig. 5).

```
Epoch [1/27], Train Loss: 0.0456, Train Acc: 0.2702, Val Loss: 0.0229, Val Acc: 0.2917
Epoch [2/27], Train Loss: 0.0447, Train Acc: 0.3636, Val Loss: 0.0227, Val Acc: 0.4500
Epoch [3/27], Train Loss: 0.0395, Train Acc: 0.5051, Val Loss: 0.0227, Val Acc: 0.4917
Epoch [4/27], Train Loss: 0.0392, Train Acc: 0.4268, Val Loss: 0.0197, Val Acc: 0.4417
Epoch [5/27], Train Loss: 0.0336, Train Acc: 0.5480, Val Loss: 0.0181, Val Acc: 0.4333
Epoch [6/27], Train Loss: 0.0291, Train Acc: 0.6086, Val Loss: 0.0208, Val Acc: 0.4083
Epoch [7/27], Train Loss: 0.0275, Train Acc: 0.6111, Val Loss: 0.0173, Val Acc: 0.5000
Epoch [8/27], Train Loss: 0.0245, Train Acc: 0.6566, Val Loss: 0.0264, Val Acc: 0.4917
Epoch [9/27], Train Loss: 0.0229, Train Acc: 0.6944, Val Loss: 0.0191, Val Acc: 0.4667
Epoch [10/27], Train Loss: 0.0210, Train Acc: 0.7247, Val Loss: 0.0179, Val Acc: 0.5083
Epoch [11/27], Train Loss: 0.0203, Train Acc: 0.7551, Val Loss: 0.0209, Val Acc: 0.4667
Epoch [12/27], Train Loss: 0.0205, Train Acc: 0.7576, Val Loss: 0.0169, Val Acc: 0.6250
Epoch [13/27], Train Loss: 0.0144, Train Acc: 0.8182, Val Loss: 0.0153, Val Acc: 0.5750
Epoch [14/27], Train Loss: 0.0142, Train Acc: 0.8106, Val Loss: 0.0214, Val Acc: 0.6500
Epoch [15/27], Train Loss: 0.0105, Train Acc: 0.8838, Val Loss: 0.0150, Val Acc: 0.5917
Epoch [16/27], Train Loss: 0.0098, Train Acc: 0.8965, Val Loss: 0.0279, Val Acc: 0.6250
Epoch [17/27], Train Loss: 0.0083, Train Acc: 0.9091, Val Loss: 0.0257, Val Acc: 0.6333
Epoch [18/27], Train Loss: 0.0082, Train Acc: 0.9040, Val Loss: 0.0126, Val Acc: 0.6417
Epoch [19/27], Train Loss: 0.0084, Train Acc: 0.9015, Val Loss: 0.0213, Val Acc: 0.6083
Epoch [20/27], Train Loss: 0.0064, Train Acc: 0.9268, Val Loss: 0.0130, Val Acc: 0.6250
Epoch [21/27], Train Loss: 0.0055, Train Acc: 0.9343, Val Loss: 0.0265, Val Acc: 0.6083
Epoch [22/27], Train Loss: 0.0061, Train Acc: 0.9116, Val Loss: 0.0290, Val Acc: 0.6583
Epoch [23/27], Train Loss: 0.0031, Train Acc: 0.9646, Val Loss: 0.0279, Val Acc: 0.5833
Epoch [24/27], Train Loss: 0.0030, Train Acc: 0.9571, Val Loss: 0.0735, Val Acc: 0.5750
Epoch [25/27], Train Loss: 0.0036, Train Acc: 0.9646, Val Loss: 0.0366, Val Acc: 0.6417
Epoch [26/27], Train Loss: 0.0027, Train Acc: 0.9823, Val Loss: 0.0336, Val Acc: 0.6500
Epoch [27/27], Train Loss: 0.0025, Train Acc: 0.9747, Val Loss: 0.0246, Val Acc: 0.6417
```

Figure 5: Disease recognition model training and validation accuracy and loss

Based on this disease recognition model, the system can automatically identify dis-ease types after inspectors upload disease images, without the need of manual judgment and manual reporting of disease types. The realization of automatic disease recognition technology can meet the efficient and accurate classification of engineering equipment diseases, and can thus provide an important data basis for the subsequent maintenance process.

4.2. Model Solution and Numerical Example of the Equipment State Judgment Model

The experimental data was collected from two high-speed railway engineering departments, and a total of 160,000 historical records of periodic inspection of engineering equipment were collected. The operating status of engineering equipment is mainly divided into in use, overhaul, update, and scrap, and there are 40,000 records in each of the four categories of status. The records of each equipment state are divided into training set, verification set, and testing set based on the ratio of 2:1:1. The information such as problem description and construction content in the inspection records can provide key basis for judging the equipment status, so the model will identify the equipment operating status by capturing such key information. The equipment state judgment model is evaluated using a macro average model, and the macro average model evaluation results are shown in Table 4.

Table 4: Equipment state judgment model evaluation results

No.	Research model	P	TPR	F1
1	CNN-LSTM	0.8903	0.9142	0.9021

All kinds of indicators have reached ideal values. Therefore, the following conclusions can be drawn: The equipment status recognition model based on CNN-LSTM can accurately judge the operating status of equipment by capturing key information in inspection records.

Based on the GM (1,1) model constructed, the predicted number of inspection times per quarter of the switch fault inspection project from January 2022 to June 2024 was calculated, and the relative error was calculated, so as to analyze the accuracy of this model. The results are shown in Table 5.

Table 5: GM (1,1) model inspection times prediction results

Year-quarter	Actual inspection times (days)	Predicted inspection times (days)	Relative error (%)
2022-1	5	6	20%
2022-2	7	8	14%
2022-3	7	9	28.60%
2022-4	8	10	25%
2023-1	11	12	9%
2023-2	10	13	30%
2023-3	15	16	6.70%
2023-4	19	18	5.30%
2024-1	21	20	4.80%
2024-2	—	23	—

The average relative error of the model is 14.3% and the accuracy is 85.7%. The average relative error is less than 20%, indicating that the accuracy of the model is desired, and the actual value has little fluctuation compared with the predicted value. Therefore, this model provides a better quantitative estimate for the number of quarterly inspection of switch fault detection projects in the future, and thus possesses important reference value.

5. High-Speed Railway Engineering Equipment Inspection and Repair Digital Implementation Case

Spring-and-autumn special inspection content varies according to the state of the line, and its inspection items include line static inspection, ballastless track structure inspection, rail disease inspection, switch rail wear inspection, and curve rail wear inspection. Inspectors need to fill in various disease information based on different inspection items. The closed-loop audit process involves multiple levels of staff, each with different using requirements. Therefore, the design of digital management for spring-and-autumn special inspection should possess the characteristics of multi-modules and multi-roles, and mainly includes four functional modules: disease entry, disease to-do list, disease done list, and image database. The specific functional structure is as follows: closed-loop process of disease recording and audit, disease entry and display of various spring-and-autumn inspection records, and systematic presentation of disease images.

The main business objective of the periodic inspection early warning function is to realize the visualization as well as the statistical analysis of the early warning information of the inspection items of engineering equipment through the combination of map and list. Users can check the inspection item information list to view the early warning status of each item and the historical inspection records associated with the item, and use the GIS map to view all the mileage range or equipment which has overdue or approaching overdue inspection items. There are three main functional modules: equipment classification card page, inspection item list page, and GIS map page.

After analyzing the key technologies of spring-and-autumn special inspection management application and periodic inspection early warning application, a high-speed railway engineering department is selected as a pilot application to explore the practical application.

5.1. Spring-and-Autumn Special Inspection Management Application

The application of spring-and-autumn special inspection management helps the inspectors to enter the disease information of various inspection items from handheld device applications and tablets (see Fig. 6), and provides a more convenient means for the on-site recording of spring and autumn diseases. Each inspection item corresponds to a label page, which improves the regularity of disease record classification.

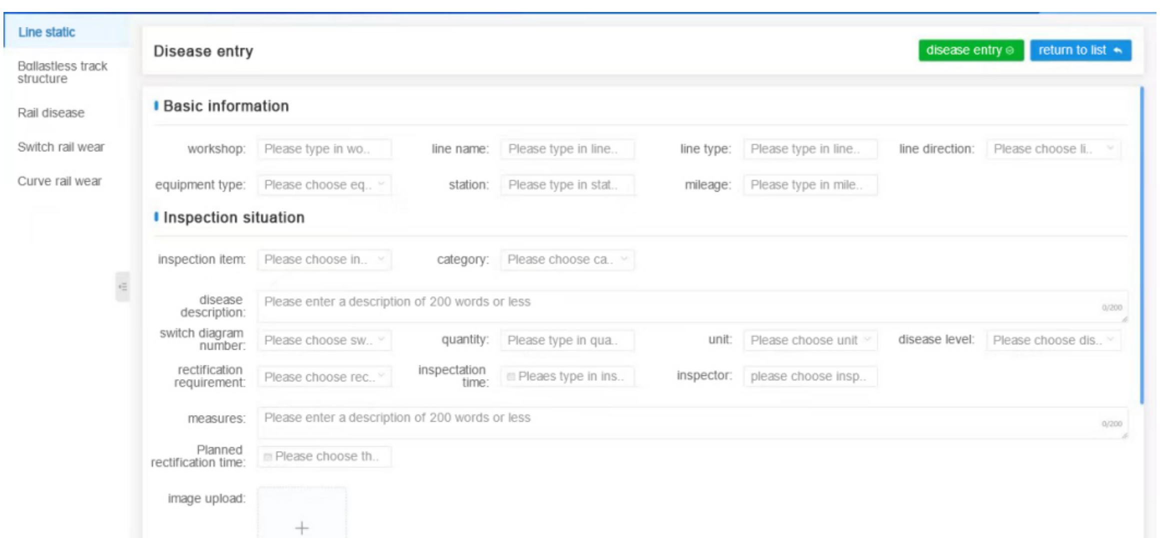

Figure 6: Spring-and-autumn special inspection management application disease entry page

Disease records are presented in the form of lists, and users of various roles can intuitively grasp disease information and the audit status of disease records. The image database module groups disease images by disease records (see Fig. 7), and each group of images is displayed in a banner. The watermark information on each image briefly summarizes the corresponding disease records, and users can intuitively obtain the correlation between the image and the corresponding disease item.

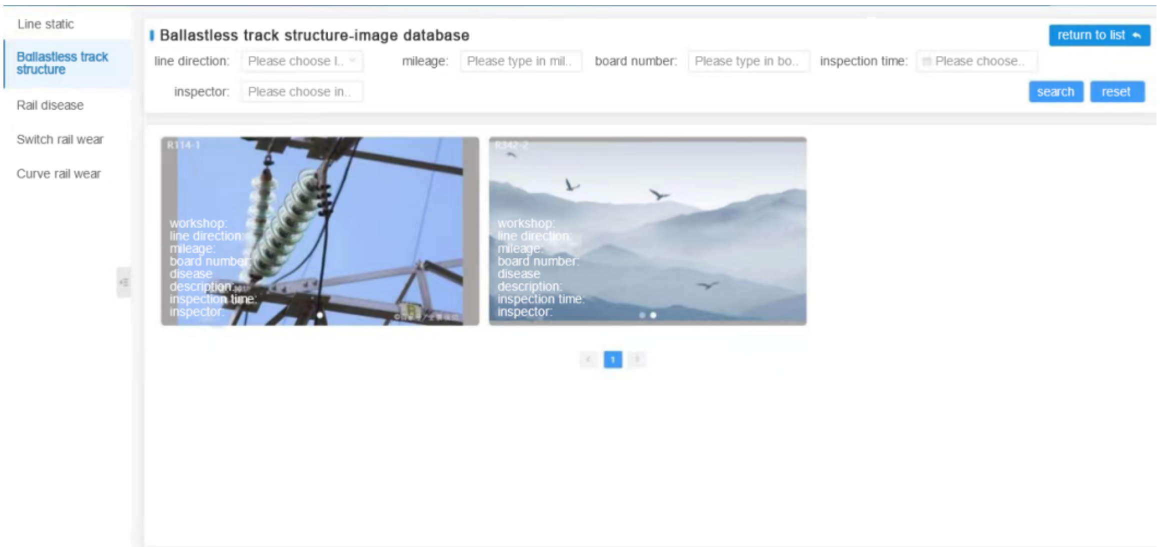

Figure 7: Spring-and-autumn special inspection management application image database page

At the same time, this application has effectively realized the closed-loop process of the record and audit of inspection disease records, realized the division of operation permission according to the role and audit status, and accelerated the closed-loop process. For example, workshop managers can fill in audit opinions on the disease audit page, and can view disease details and historical audit information at the same time. The integrated page design of audit information filling and detailed information helps to improve audit efficiency.

Finally, the application of spring-and-autumn special inspection management has also realized automatic disease recognition. After the inspector uploads the disease picture in the disease record module, the application will automatically identify the disease category, and fill the identification result into the ”disease category” field without the need of manual input. Compared with the traditional manual judgment and filling mode, this application has achieved higher accuracy and efficiency.

In summary, the application of spring-and-autumn special inspection management has broken the limitations of traditional manual recording of disease information, realized high quality, high visualization, and high efficiency of disease recording, and provided technical support for systematic disease record audit, which can effectively help inspectors and managers of engineering departments to improve the overall efficiency of spring and autumn inspection and maintenance.

5.2. Periodic Inspection Early Warning Application

Early warning status and the number of early warning items of various kinds of equipment are displayed on the equipment classification card, inspection item list page, and GIS map page. The home page integrates warning status and statistical information in an intuitive and eye-catching way (see

Fig. 8). The color of each equipment classification card represents the overall inspection warning status of this equipment type, and the number of warnings of the equipment type is displayed on each equipment card. Select an equipment category card from the home page to enter the inspection item list page of the corresponding type of equipment. A list page consists of multiple tabs, each representing a type of inspection item under the equipment category. The list page provides two ways to view the early warning status of inspection items: the color identifier in the serial number column represents the early warning status, and the inspection items are also sorted according to the early warning status (see Fig. 9). The overdue items are ranked first to remind users that these inspection items need to be paid great attention to; Fields such as Remaining Days and Warning Status in the list provide text information for determining the warning status. In addition, this page displays the number of overdue and approaching overdue inspections for each type of inspection item by equipment type in the form of a "number + color" identifier.

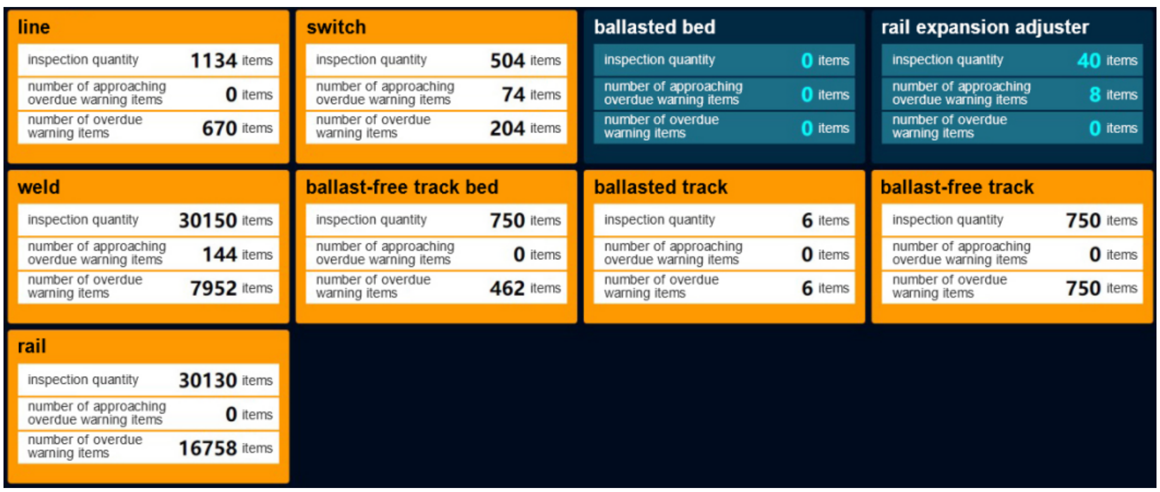

Figure 8: Periodic inspection early warning equipment classification card page

Figure 9: Periodic inspection early warning list page

The GIS map page also integrates warning status and quantitative statistics. This page presents the warning status by color-identified sections. Users can combine the color identifier with the warning information pop-up window to view the inspection warning status of all equipment in a specified mileage range. The number of warnings is displayed on the equipment type tab page, including the number of overdue and approaching overdue inspections for each type of equipment.

In summary, the periodic inspection early warning application has realized the intuitive and eye-catching presentation of the early warning status and early warning statistical information of each inspection item. Through the combination of graph and table, color identification and text information, on-site users can obtain information about which equipment has overdue inspection items and which equipment type has the most serious overdue inspection. Therefore, this application design provides sufficient technical support for field workers to carry out inspection and repair analysis and decision efficiently and accurately.

6. Conclusion

Based on the comprehensive analysis of the current situation of inspection and repair management of high-speed railway engineering equipment, this paper designs the application of special maintenance management in spring and autumn and the application of periodic inspection early warning, researches the key technologies such as deep learning and big data, and establishes the automatic disease recognition model, equipment status judgment model, and recommended inspection date prediction mechanism. Through the pilot application of the system, this paper fully proves that the design and fusion of the two applications can provide sufficient technical support for the inspection and repair of high-speed railway engineering equipment, and can provide a comprehensive and reliable data basis for real-time and efficient maintenance decisions. In the next step, more technical models and data dimensions will be integrated to support more accurate decision-making, and continue to enhance the security of high-speed rail transportation.

References

Yanguo Huang, Chencong Zhou, and Kefei Zuo. Dlstm-ae traffic flow prediction model based on radam optimization. *Highway traffic technology*, 40(1):185–191,199, 1 2023.

Yixin Huang, Wenshan Fang, and Chuanpeng Liu. House extraction from railway remote sensing images based on improved mask r-cnn model. *Railway computer application*, 33:7–12, 2024. ISSN 1005-8451.

Yunqing Huang, Leilei Wang, Mingze Xu, Peizheng Hou, and Haifeng Yan. Redundant vertex optimization of 3d printing model based on hash table. *Technology and market*, 29(11):29–32, 11 2022.

Md Rakibul Islam, Golam Kabir, Kelvin Tsun Wai Ng, and Syed Mithun Ali. Yard waste prediction from estimated municipal solid waste using the grey theory to achieve a zero-waste strategy. *Environmental Science and Pollution Research*, 29(31):46859–46874, 2022. ISSN 1614-7499. doi: 10.1007/s11356-022-19178-y.

K, Subhashini, R, Aarthi Lakshmi, V, Arthi, and G, Hemalatha. Image encryption using convolutional neural network. *ITM Web Conf.*, 56:05005, 2023. doi: 10.1051/itmconf/20235605005.

Honggang Pan, Yuanlu LI, Baoren Guo, and Yubo Liu. Research and diagnosis of rotor system faults based on convolutional neural networks. *Turbine technology*, 66(2):145–148, 4 2024.

Junli Peng, Yu Gu, Zheng Zhang, and Xiaohan Geng. Document representation fused with term contribution and word2vec word vector. *Computer engineering*, 47:62–67, 2021. ISSN 1000-3428. doi: 10.19678/j.issn.1000-3428.0056370.

Jizhou Ren. Job posting classification research based on deep learning. *Journal of Hubei University of Arts and Sciences*, 44(11):21–27, 11 2023.

Junqing Shi, Linwu Chen, Sulan Li, Guolian Meng, and Shunya Xia. Urban road traffic incident detection algorithm based on cnn. *Highway traffic technology*, 39(3):176–182, 3 2022.

Tianyun Shi. Present situation of wide applications of information and intelligence in the field of high-speed railway in china. *Science and technology guide*, 37:53–59, 2019. ISSN 1000-7857.

Joongbo Shin, Yanghoon Kim, Seunghyun Yoon, and Kyomin Jung. Contextual-cnn: A novel architecture capturing unified meaning for sentence classification. In *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 491–494, 2018. doi: 10.1109/BigComp.2018.00079.

Ying Wang, Zhengqiu Liu, Chengshuai LI, Jinlin Bai, Wenjie Yang, and Yong Zhao. Mine water inflow prediction based on armi-gm model. *Coal technology*, 2024.

Qingheng Xie. Analysis of several influencing factors on word2vec text classification effect. *Modern information technology*, 8:125–129, 2024. ISSN 2096-4706. doi: 10.19850/j.cnki.2096-4706.2024.01.026.

Huaizhi Yang, Yanhua Wu, and Zhibo Cheng. Big data application of high-speed railway track maintenance equipment based on railway data service platform. *Railway computer application*, 27(7):85–89, 1 2018.

Xin Zhang, Zhengli Zhai, and Luyao Yao. Text classification of chinese news based on cnn and lstm hybrid model. *Computer and digital engineering*, 51(7):1540–1543,1573, 7 2023.

Jinzi Zheng, Rui Xue, Yanhua Wu, and Ping Li. Big data research and application of international railways. *Chinese railway*, pages 54–62, 2018. ISSN 1001-683X. doi: 10.19549/j.issn.1001-683x.2018.02.054.