# A Simple Heterogeneous Transfer Learning Method for Track Circuit Fault Prediction

Proc IMechE Part F: J Rail and Rapid Transit
XX(X):1–13

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

**Lan Na[1] Baigen Cai[1,2] Chongzhen Zhang[3] Jiang Liu[1,2] and Zhengjiao Li[1,2]**

**Abstract**

Prediction and identification of faults in track circuit are crucial for improving the safety and efficiency of railway transportation. However, the task of track circuit fault prediction through deep learning methods facing significant challenges due to the absence of reliable data. In this paper, a novel heterogeneous transfer learning method is proposed, aiming to reduce track circuit data reliance in model training by using publicly available datasets in other similar fields. An index describing the data distribution based on autoencoder feature extraction and maximum mean discrepancy is used to demonstrate the transferability between heterogeneous data firstly. Then a heterogeneous transfer learning method is constructed to accelerate track circuit fault prediction model training. Furthermore, the resulting deep learning model is compared to existing fault prediction methods. Finally, by adjusting the degree of involvement of transfer learning throughout model training, this paper comprehensively examines its effect on model training process. The simulation experimental results show that the proposed method can transfer useful knowledge in other similar fields for tasks in track circuit fault prediction, and the resulting model can correctly classify over 99% on the test dataset while reducing the amount of required track circuit data to 10% of the traditional training methods. The relevant methods proposed in this paper can significantly enhance the practical application value of fault prediction models based on deep learning methods in the field of intelligent maintenance of track circuit.

**Keywords**

Track circuit, fault prediction, fault diagnosis, heterogeneous transfer learning, domain adaptation

## Introduction

Track circuit equipment is essential for ensuring transportation safety in the railway system. It plays a crucial role in the smooth running of trains and helps prevent accidents.[1] Therefore, it is crucial to maintain and update track circuit equipment regularly to meet safety standards. However, track circuit equipment operates in demanding conditions, including shocks and vibrations. Electrical and physical damage arising from such poses a significant threat to reliable and secure railway operations.[2]

In order to ensure the safety and optimal functionality of the railway network, preventing track circuit faults is of utmost importance. Currently, track circuit maintenances are predominantly carried out through human experience, which is less efficient. To enhance the efficiency of track circuit maintenances, researchers have increasingly turned to machine learning methods for intelligent fault diagnosis in recent years.[3–5] The utilization of machine learning algorithms has significantly improved track circuit fault diagnosis efficiency, making it a viable alternative approach to achieving intelligent track circuit maintenances.

The approach of track circuit fault prediction is using Deep Neural Network (DNN) which delivered a new development prospect of Intelligent Maintenance (IM) research for track circuit. Theoretically, a Deep Learning (DL) model can finely fit any complex high-dimensional nonlinear function, allowing for flexible solutions to fault prediction tasks. Kang et al. put forward a DL model for predicting the number of compensation capacitors needed for track circuit maintenance using Long Short-Term Memory network

[1]School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing, China
[2]Beijing Engineering Research Centre of EMC and GNSS Technology for Rail Transportation, Beijing, China
[3]Shuohuang Railway Development Co.,Ltd., Hebei, China

**Corresponding author:**
Baigen Cai, School of Electronics and Information Engineering, Beijing Jiaotong University, Haidian, Beijing, 100044, China.
Email: bgcai@bjtu.edu.cn

(LSTM).[6] Shi et al. make research on switch equipment fault prediction method based on Gated Recurrent Unit (GRU).[7] Dai et al. make the fault prediction study based on data provided by a major rail transit agency in the United States that can identified around one-third of signal failures one month in advance by concentrating on 10% of locations on the network.[8]

Although DL models can achieve fault prediction of track circuit, it usually necessitates a substantial quantity of labelled data for training, which frequently requires a lengthy period and extensive resources for the collection and labelling of data. Especially for researches on track circuit fault prediction, researchers can hardly collect enough real data for DL model training. Most researchers have attempted to reduce the cost of data collection by using data augmentation algorithms or software simulation.[9,10] However, data obtained through these techniques fails to satisfy the crucial basic assumptions of classical machine learning as they do not achieve independent and identical distribution with the actual data.[11] Consequently, the generalization capability of DL models is reduced in practice, even leading to total malfunctioning in some instances.[12,13] And a few other researchers acquire data through performing laboratory experiments, yet this approach proves to be exceedingly expensive.[14]

In order to address the issue of insufficient training data for DL models aimed at track circuit fault prediction, the primary solution is to implement transfer learning. Transfer learning can decrease the requirement of track circuit data for track circuit fault prediction DL model training by acquiring empirical knowledge from related tasks. Many applications of homogeneous transfer learning can be found in the literature especially in the field of bearing IM. Tang et al. put forward a novel lightweight transfer learning network is proposed that can adaptively select the input length and accurately identify the bearing health states under different work conditions.[15] An other example given in is that Su et al. proposed a convolution deep belief network-dynamic multilayer perceptron for bearing fault recognition under alterable running states.[16]

Not like IM researches on bearing or other fields, in the field of railway track circuit fault prediction researches lies in the lack of public datasets available to researchers which makes a great challenge for DL researches in track circuit fault prediction. And this problem can be solved through introducing heterogeneous transfer learning that using public datasets in other relevant research fields to help track circuit fault prediction DL model training.

There are three major contributions of this paper:

1) A track circuit fault prediction DL model based on Bidirectional LSTM (Bi-LSTM) and multi-head attention mechanism is proposed which introduce multi-head attention mechanism into track circuit fault prediction innovatively. The proposed model can extract track circuit fault omen features and output the prediction result in real time. And it is able to overcome the difficulty of extract high-dimensional time series data features which is able to achieve track circuit fault prediction with high accuracy and efficiency.

2) A comprehensive index for heterogeneous transfer learning assessment is proposed. This index can quantitative analysis the DL model transferability between different domains from the aspect of data distribution that can facilitate the evaluation of transfer learning's feasibility before its implementation.

3) A novel heterogeneous transfer learning method is proposed to utilize publicly available aircraft engine degradation dataset to reduce the track circuit fault prediction DL model training data requirements so that enhance the application value of DL methodology in practical track circuit maintenance work and thereby raise the efficiency of track circuit maintenance.

The overall framework of the heterogeneous transfer learning-based track circuit fault prediction method proposed in this paper is shown in Figure 1.

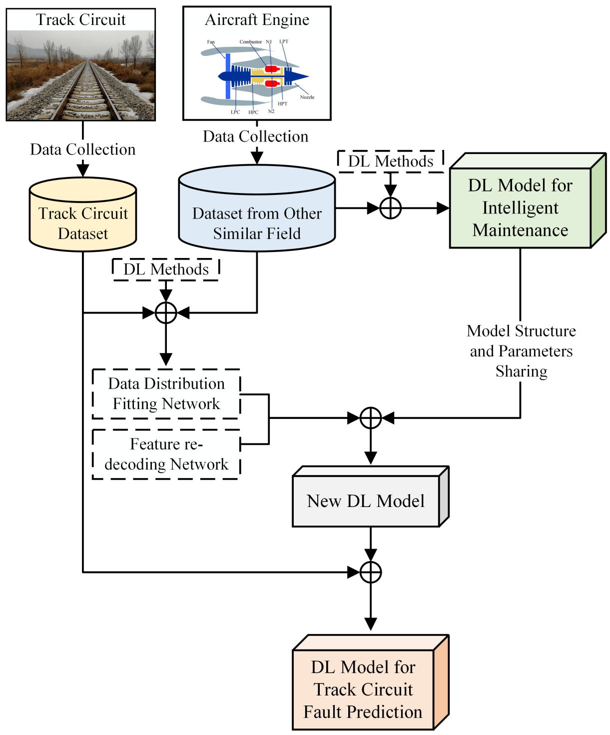


**Figure 1.** Overall framework of proposed track circuit fault prediction method.

## IM methods based on DNN

### General structure of DL-based IM methods

The architecture of IM methods based on DL can be summarized as Figure 2 which is a typical kind of encoder-decoder structure. The $\boldsymbol{x}$ refers to input data, such as data from track circuit or aircraft engine monitoring sensors. The function of input layer is normalize input data $\boldsymbol{x}$ to enhance the training speed and convergence progress of the DL model.

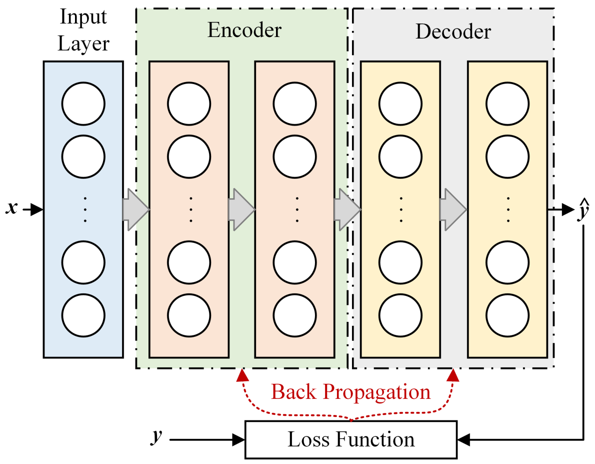


**Figure 2.** General IM methods structure based on DL.

The encoder's role is to extract high-dimensional features from the input data. And encoder can be multi-layers so that it can extract more complex high-dimensional features. The most commonly used component units for encoder including fully connected layer, convolutional layer, recurrent neural layer, and etc. The output of encoder i.e., the high-dimensional features of input data, can be seen as the key information describing the degradation process of an equipment.

The decoder's role is to map the high-dimensional data features produced by the encoder onto an expected output as $\hat{\boldsymbol{y}}$ in this figure. Then the loss function can calculate model loss by $\hat{\boldsymbol{y}}$ and real label $\boldsymbol{y}$, and then the model parameters can be updated by backpropagation algorithm.

### DL-based prognostics methods

Prognostics is an important research area in the field of IM, and its main research includes fault prediction and Remaining Useful Life (RUL) prediction. In the researches of prognostics, researchers commonly employ DL methods like Recurrent Neural Network (RNN) and attention mechanism.[17,18] These DL methods can effectively mine the time serial features of data information to achieve prognostics. The following describes DL methods and general DL model structure used in prognostics study.

For the tasks of prognostics, LSTM is a logical option as its recurrent connections enable the storage of past events. LSTM networks have the capability to learn long-term time dependences by incorporate specialized memory cells into the network architecture.[19] The structure of LSTM unit is shown in Figure 3.

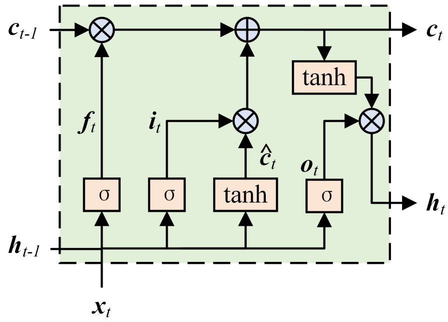


**Figure 3.** LSTM unit structure.

At the current time step $t$, the input to each LSTM cell consists of the input data at that time step $\boldsymbol{x}_t$, and the outputs of all LSTM cells at the previous time step $\boldsymbol{h}_{t-1}$. The equations that describe the relationships between input and output of LSTM unit are:

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W}_{xf} \otimes \boldsymbol{x}_t + \boldsymbol{W}_{hf} \otimes \boldsymbol{h}_{t-1} + \boldsymbol{b}_f) \tag{1}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_{xi} \otimes \boldsymbol{x}_t + \boldsymbol{W}_{hi} \otimes \boldsymbol{h}_{t-1} + \boldsymbol{b}_i) \tag{2}$$

$$\hat{\boldsymbol{c}}_t = \tanh(\boldsymbol{W}_{xc} \otimes \boldsymbol{x}_t + \boldsymbol{W}_{hc} \otimes \boldsymbol{h}_{t-1} + \boldsymbol{b}_c) \tag{3}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_{xo} \otimes \boldsymbol{x}_t + \boldsymbol{W}_{ho} \otimes \boldsymbol{h}_{t-1} + \boldsymbol{b}_o) \tag{4}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \otimes \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \otimes \hat{\boldsymbol{c}}_t \tag{5}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh(\boldsymbol{c}_t) \tag{6}$$

while $\otimes$ in these equations means Hadamard product with two vectors.

In order to excavate the contextual semantic information in the current time step, this paper introduces bidirectional structure in the LSTM to obtain the contextual feature.[20] Figure 4 shows the structure of the Bi-LSTM unfolding in three time steps.

The Bi-LSTM network contains forward and backward network. The $\boldsymbol{h}_t^f$ and $\boldsymbol{h}_t^b$ are the output of forward and backward network at current time step. And the output of Bi-LSTM at time step $t$ can be written as $\boldsymbol{h}_t = [\boldsymbol{h}_t^f; \boldsymbol{h}_t^b]$.

And to further enhance the performance of the model while mitigating the overfitting problem, this paper adds multi-head attention mechanism[21,22] behind the Bi-LSTM structure. The essence of the multi-head method is to perform multiple independent attention mechanism calculations, and

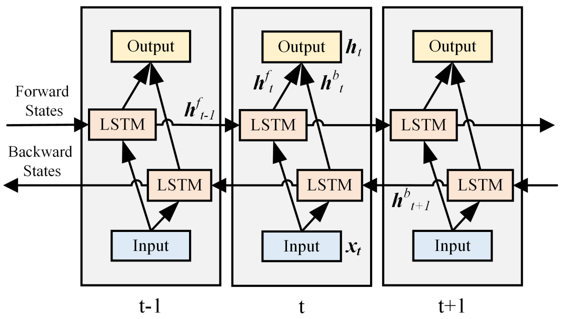


**Figure 4.** Structure of bidirectional LSTM.

splicing the calculation results into an integrated result. The structure of multi-head attention mechanism is shown in Figure 5. And the equations that describe the input and output of relationships between multi-head attention mechanism are:

$$\boldsymbol{m}_i = \tanh(\boldsymbol{W}_i^m \otimes \boldsymbol{x}_i) \tag{7}$$

$$\boldsymbol{s}_i = \mathrm{Softmax}(\boldsymbol{m}_i) \tag{8}$$

$$z_j = \sum_{i=1}^{n} \boldsymbol{x}_i \otimes \boldsymbol{s}_i \tag{9}$$

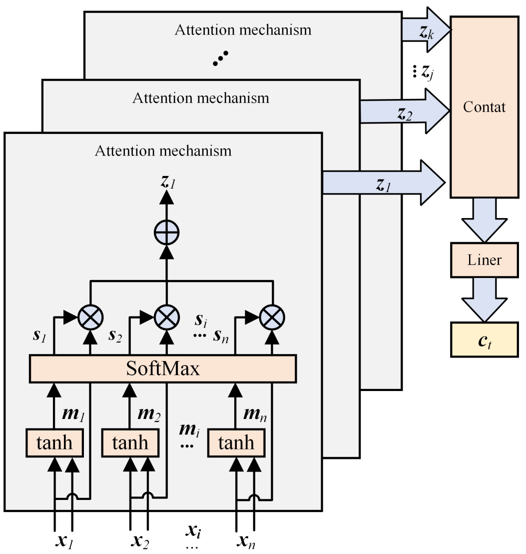


**Figure 5.** Structure of multi-head attention mechanism.

The output form after Contat is $[z_1; z_2; ...; z_k]$, and the Linear layer is a single-layer fully-connected neural network without nonlinear activation. The $\boldsymbol{c}_t$ is the final output.

So, the structure overview of general prognostics DL model can be summarized as Figure 6. There are multi layers of Bi-LSTM containing several LSTM cells each, followed by one multi-head attention layer with several heads. Then followed by Flatten & Dropout structure.

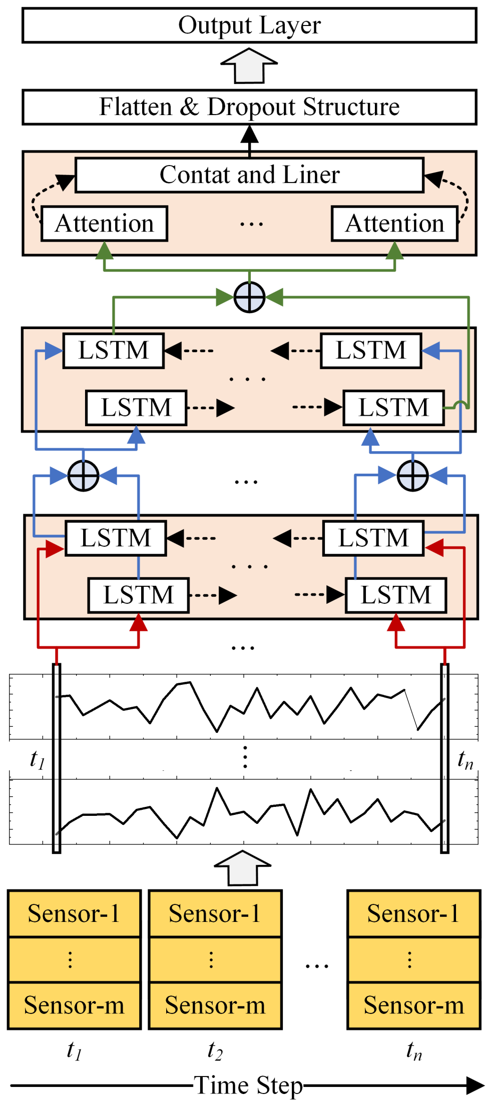


**Figure 6.** Structure of general prognostics DL model.

## Transfer learning method for track circuit fault prediction

### *Heterogeneous transfer learning*

Heterogeneous transfer learning involves utilising knowledge from the source domain to enhance model performance in the target domain. And the heterogeneous approach assumes distinct feature spaces, data distributions and labelling spaces between the source and target domains. This assumption is more practical in real-world applications.[23] For example, this paper attempts to using aircraft engine RUL prediction data to help achieve track circuit fault prediction since there is a shortage of research data in the field of track circuit fault prediction, whereas there are large publicly datasets in the field of aircraft engine RUL prediction.

The issue with heterogeneous transfer learning lies in how to address dissimilarities between the source and target domains. The commonly used heterogeneous transfer learning approaches include domain adaptation, multiple instance learning, heterogeneous feature selection, heterogeneous feature mapping, and heterogeneous label transformation.[24] Since there are differences in both the data and label between aircraft RUL prediction and track circuit fault prediction, a combination of heterogeneous feature mapping and heterogeneous label transformation methods is necessary for effective heterogeneous transfer learning.

### Transfer learning assessment index

It is clear that there is no direct link between aircraft engine RUL prediction and track circuit fault prediction. In general, however, both tasks are concerned with the process of degradation of a particular piece of equipment. In order to show the DL model transferability between two tasks, an evaluation index is proposed to assess the difference in the distribution of the data between the two tasks.

Autoencoder is an unsupervised artificial neural network that utilises encoder-decoder structure. The encoder compresses high-dimensional data into lower-dimensional data (the code), while the decoder then attempts to reconstruct the original high-dimensional data.[25] The ultimate goal of this process is to obtain a representation of the data in lower dimensions. So the transfer learning assessment index can be constructed through autoencoder uniform dimensional extraction of data features. The structure of the autoencoder based feature extractor is shown in Figure 7.

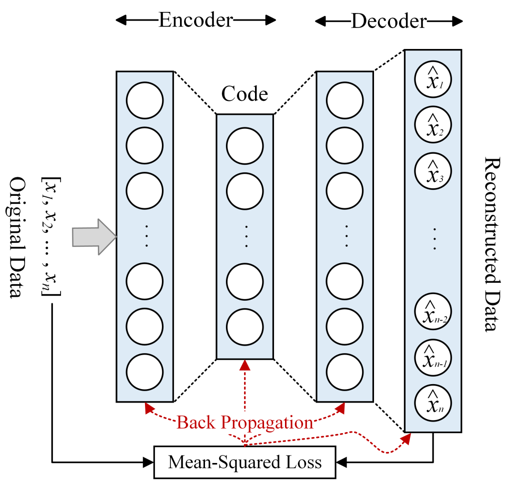


**Figure 7.** Structure of autoencoder based feature extractor.

After feature extracted by using autoencoder, the transfer learning evaluation index is constructed using the MMD loss. To calculate the MMD loss $\hat{d}_{mmd}$, refer to equation (10) where $\boldsymbol{o}^c_\cdot$ and $\boldsymbol{o}^a_\cdot$ represent feature extraction outcomes of the data from the two tasks, respectively. The $n_c$ and $n_a$ represent the number of sample data for two tasks, respectively. And $k(\boldsymbol{a},\boldsymbol{b}) = \Phi(\boldsymbol{a}) \cdot \Phi(\boldsymbol{b})$, which $\Phi(\cdot)$ means the feature map to map the data $\boldsymbol{a}$ and $\boldsymbol{b}$ to the reproduction kernel Hilbert space. The result of $\hat{d}_{mmd}$ is using as the transfer learning assessment index.[26]

$$\hat{d}_{mmd} = \frac{1}{n_c^2}\sum_{i=1}^{n_c}\sum_{j=1}^{n_c} k(\boldsymbol{o}_i^c, \boldsymbol{o}_j^c) + \frac{1}{n_a^2}\sum_{i=1}^{n_a}\sum_{j=1}^{n_a} k(\boldsymbol{o}_i^a, \boldsymbol{o}_j^a) - \frac{1}{n_c n_a}\sum_{i=1}^{n_c}\sum_{j=1}^{n_a} k(\boldsymbol{o}_i^c, \boldsymbol{o}_j^a) \quad (10)$$

### Transfer learning-based model training method

In order to apply the knowledge gained from aircraft engine RUL prediction to facilitate fault prediction of track circuit aiming to reduce the data dependency within the model training process, this paper proposes a novel heterogeneous transfer learning method for track circuit fault prediction DL model training and its structure is shown in Figure 8. The structure of trained aircraft engine RUL prediction model in this figure is designed as Figure 6, which has two Bi-LSTM layers containing 128 LSTM cells each, one multi-head attention mechanism layer with eight heads and two dropout layers with dropout probability 0.25 and 0.55 separately. And the model output layer has one unit to output the RUL prediction result.

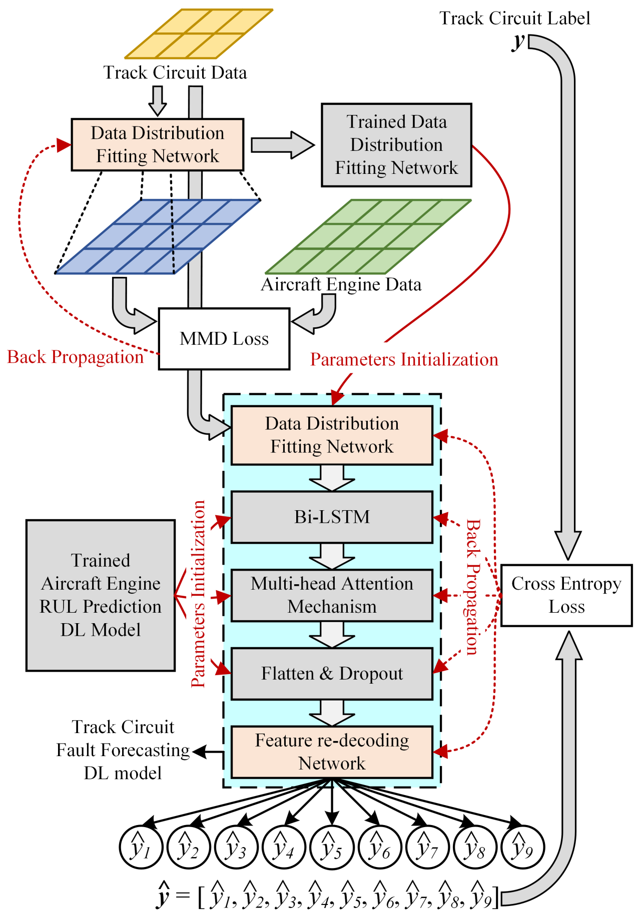


**Figure 8.** Transfer learning-based model training method.

The key components in this method are the data distribution fitting network and the feature re-decoding network. The two networks correspond to heterogeneous feature mapping and heterogeneous label transformation methods, respectively. The role of the data distribution fitting network is to harmonise the dimension and distribution of the track circuit data with the aircraft engine data, and its specific structure is shown in Figure 9.

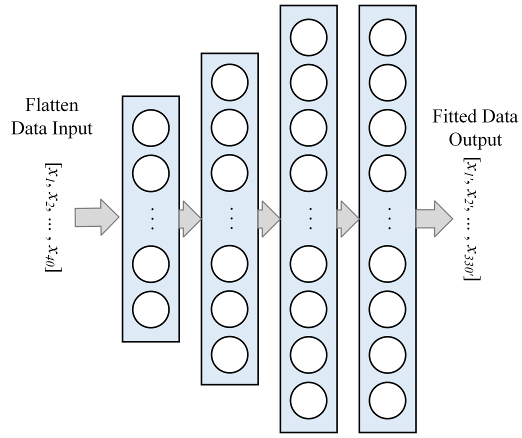


**Figure 9.** Structure of data distribution fitting network.

The data distribution fitting network contains 4 fully connected layers of 110, 220, 330 and 330 units respectively. The purpose of designing the data distribution fitting network as multiple fully connected layers is to prepare for better fit the data distribution of aircraft engine RUL prediction data and track circuit fault prediction data.

And the specific structure of feature re-decoding network is similar with the structure of data distribution fitting network which is shown in Figure 10. The feature re-decoding network contains 3 fully connected layers of 2048, 1024 and 9 units respectively.

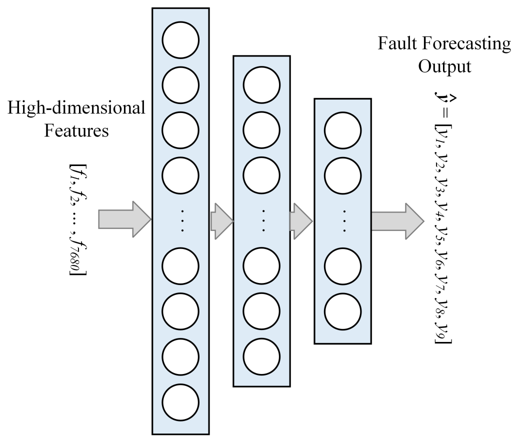


**Figure 10.** Structure of feature re-decoding network.

The feature re-decoding network within the heterogeneous transfer learning method is notably more intricate than the decoder component of the aircraft engine RUL prediction DL model, i.e., the one unit in output layer. This is due to the need to consider the significant heterogeneous distinctions in output between the two relevant tasks which makes the original decoder component lacks the capability of mapping the extracted high-dimensional feature information from the encoder to track circuit fault prediction output space. Therefore, a more complex feature decoding networks need to be designed so as to map the high-dimensional data features extracted by the model to the labelling space of the track circuit fault prediction task.

The process of implementing heterogeneous transfer learning-based model training can be summarized as Figure 11.

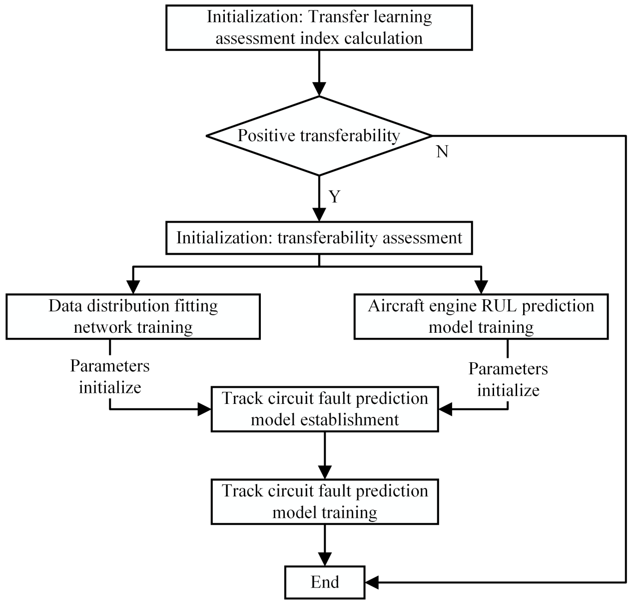


**Figure 11.** Flowchart of heterogeneous transfer learning-based model training method.

(1) The initial phase of heterogeneous transfer learning is to examine the transferability from aircraft engine RUL prediction task to track circuit fault prediction task by using the proposed transfer learning assessment index.

(2) If the transferability between two tasks is acceptable, the data distribution fitting network can be trained by using track circuit data and aircraft engine data with a loss function defined as MMD. And the aircraft engine RUL prediction model can be trained by relevant dataset.

(3) Track circuit fault prediction model is then established as depicted in blue background frame in Figure 8. The parameters of relevant layers are initialised using the corresponding parameters present in the trained aircraft engine RUL prediction model and trained data distribution fitting network.

(4) The constructed network, developed in processes (3), can then be trained utilising track circuit data and labels.

And the cross-entropy loss will be implemented as the loss function.

(5) The heterogeneous transfer learning-based training process is finally achieved, the resulting model can predict the classes of fault that will occur in the future based on the track circuit data. And the knowledge gained from the aircraft engine RUL prediction task can be used to significantly speed up the training process of the track fault prediction task and reduce the amount of training data required for the track fault prediction task.

## Experiments and analysis

This section tries to present the important results of the current works. To verify the experiment results in this paper, the experimental data of C-MAPSS aircraft engine RUL prediction dataset[27] (FD001), and track circuit fault prediction dataset[28] (TC-A and TC-B) had selected to validate this study. The function of FD001 dataset is train the aircraft engine RUL prediction model, aiming to gain the relevant knowledge so that can use the proposed transfer learning method to help track circuit fault prediction model training. The details of datasets used in this paper is shown in Table 1.

**Table 1.** Details of used datasets.

| Dataset name | Size of dataset | Number of classes |
|---|---|---|
| FD001 training | 14241 | - |
| FD001 test | 100 | - |
| TC-A training | 90 | 9 |
| TC-A test | 810 | 9 |
| TC-B | 225 | 9 |

| Class code | Class label |
|---|---|
| F0 | No failure |
| F1 | Sender failure |
| F2 | Sender cable failure |
| F3 | Track ballast resistance failure |
| F4 | Compensation capacitor failure |
| F5 | Receiver cable failure |
| F6 | Attenuation redundancy controller failure |
| F7 | Sender transformer failure |
| F8 | Receiver transformer failure |

### *Transferability assessment*

Firstly, assessment of the transfer learning is carried out to verify the transferability performance of the proposed modelling solution. The encoder of the feature extractor has two layers with 128 and 10 units, respectively. So, the size of extracted features are $1 \times 10$ vectors. And the decoder of the feature extractor has two layers with 128 and $n$ units, respectively. The value of $n$ depends on the data dimension of the original input data. The learning rate is set at 0.001, and batch size is 16. And the Mean Square Error (MSE) shown in equation 11 will be implemented as the loss function. And $\hat{\boldsymbol{x}}_i$ is the reconstructed data, $\boldsymbol{x}_i$ is the original data.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i)^2 \tag{11}$$

The heatmap of the extracted features correlation is shown in Figure 12. Figure 12 corresponds to the extracted features of the TC-A data. And Figure 13 corresponds to the extracted features of the FD001 data.

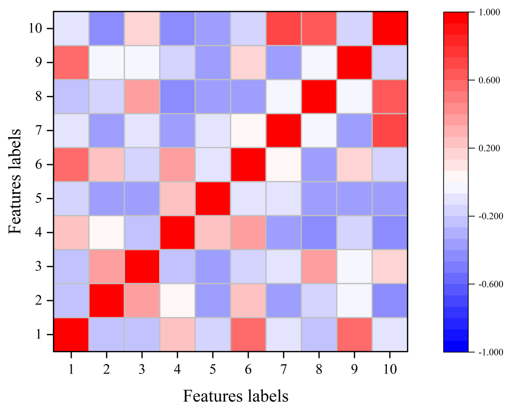


**Figure 12.** Correlation heatmap of TC-A extracted features.

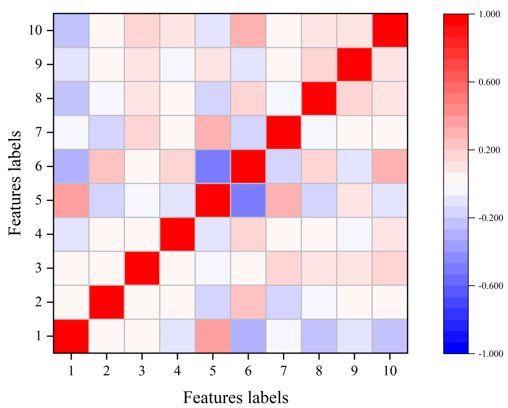


**Figure 13.** Correlation heatmap of FD001 extracted features.

As depicted in Figure 12 and Figure 13, the correlation measure illustrates a less significant linear relationship between the 10 extracted features from the original data, which can also be a side effect to demonstrate that the extracted features might possess a strong non-linear relationship with each other. Then the transfer learning assessment index is calculated from FD001 to TC-A by using equation (10).

The transfer learning assessment index value shows the absolute distance of the probability distribution between two datasets. Yet in the absence of an appropriate comparative benchmark, a isolated index will not aid in determining the level of complexity regarding transfer learning between the two datasets. To address this issue, this paper undertakes the uniform approach to compute the transfer learning assessment index from TC-A to TC-B. These two datasets were collected from identical type of track circuit situated in distinct environments, and the solitary discrepancy between them is the variation in data distribution. As such, the transfer learning assessment index, from TC-B to TC-A, can serve as a suitable benchmark to be compared.

In order to compute the transfer learning assessment index from TC-B to TC-A, extract the TC-B features using the same feature extractor initially. The heatmap of the extracted features of the TC-B data is shown in Figure 14.

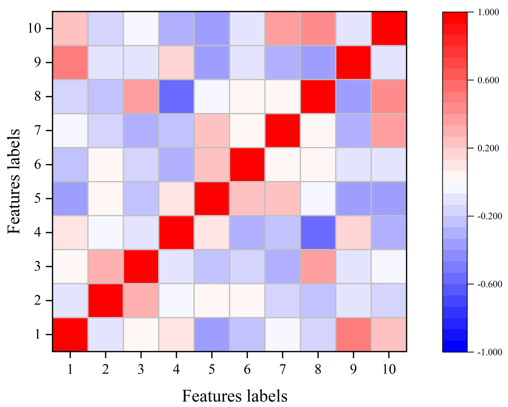


**Figure 14.** Correlation heatmap of TC-B extracted features.

To further evaluate the efficiency of the extracted features, the Root Mean Square Error (RMSE) is then computed to determine the deviation between the original data and the reconstructed data as shown in equation 12.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i)^2} \tag{12}$$

The RMSE results are summarized in Table 2. And it can be seen that the extracted features reflect the original data's characteristics effectively.

**Table 2.** The values of RMSE.

| Original dataset name | RMSE value |
|---|---|
| TC-A | 0.26 |
| TC-B | 0.21 |
| FD001 | 0.08 |

Then calculate the transfer learning assessment index, from TC-B to TC-A, using equation (10). And the two transfer learning assessment indices mentioned above are summarised in Table 3.

**Table 3.** The value of transfer learning assessment index.

| Transfer direction | Index value |
|---|---|
| FD001 to TC-A | 2.1618 |
| TC-B to TC-A | 2.0459 |

The assessment index for transfer learning from FD001 to TC-A is slightly higher compared to that of TC-B to TC-A, as shown in Table 3. Therefore, we can infer that the transferability from FD001 to TC-A is similar to that from TC-B to TC-A.

## *Aircraft engine RUL prediction*

Based on demonstrating the transferability of aircraft engine and track circuit data, this paper conducted further heterogeneous transfer learning experiments. Firstly, the aircraft engine RUL prediction DL model is constructed, and train the model by utilising the FD001 training dataset. The learning rate of training is set at 0.001, and without using mini-batch in training process. The MSE loss of the DL model in each training epoch is shown in Figure 15.

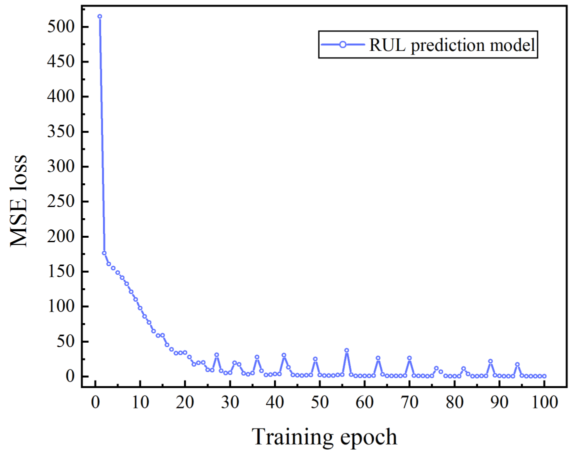


**Figure 15.** Aircraft engine RUL prediction DL model training process.

To further evaluate the RUL prediction performance of the trained DL model, this paper compared the RUL prediction Mean Absolute Error (MAE) in FD001 test dataset with other RUL prediction methods. The calculation method of MAE is shown in equation 13. In this equation, $T_i$ is the real RUL value, $\hat{T}_i$ is the predicted RUL value and $n_i$ is the current useful life.

$$MAE = \frac{|\hat{T}_i - T_i|}{n_i + T_i} \tag{13}$$

The MAE values of the aircraft engine RUL prediction results with different methods at different actual RUL levels is shown in Figure 16.

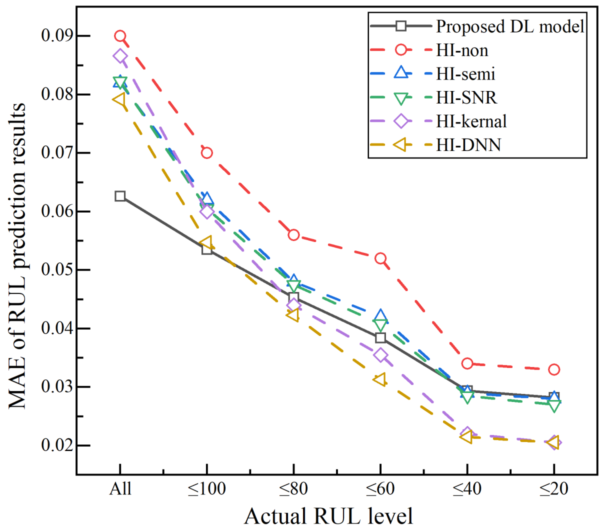


**Figure 16.** MAE values on aircraft engine RUL prediction results with different methods.

In this figure, the dash lines are the prediction errors of each different existing aircraft engine RUL prediction methods, i.e., the nonparametric method[29], the semiparametric method[30], the SNR-based method[31], Kernel-based method[32], and DNN-based method[14]. The solid line is the MAE value plotted based on the prediction results of our proposed method. And the MAE variances of the RUL prediction results for different methods is shown in Figure 17.

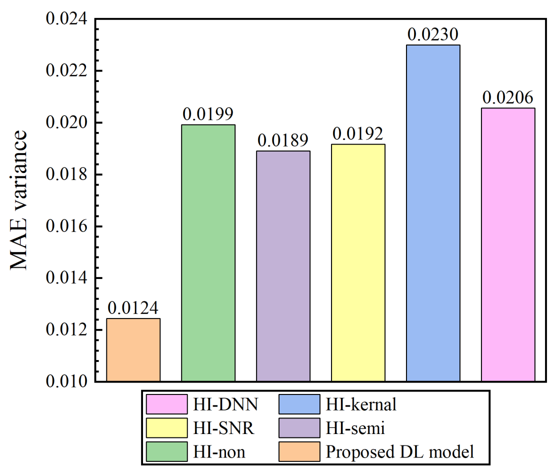


**Figure 17.** MAE variances of RUL prediction results with different methods.

It can be observed that the mean of the RUL prediction result MAE by proposed DL model achieves the minimum error in general with a significance prediction improvement in the early stage of actual RUL levels than other methods, which means that the proposed method is an available RUL prediction method. Moreover, the variance of the RUL prediction results MAE by proposed DL model is also very small, thereby indicating a stable RUL prediction of the proposed method.

### *Track circuit fault prediction*

The training process of data distribution fitting network is carried out by using FD001 and TC-A data, and the heterogeneous transfer learning model training method is constructed as shown in Figure 8. Then, the TC-A training dataset is used to train the DL model for track circuit fault prediction by the heterogeneous transfer learning-based method.

To preliminary confirm whether the knowledge gained from DL model learning in the aircraft engine RUL prediction task can hasten the training of the track circuit fault prediction model, this paper compared the aforementioned training process and the training process of same structural track circuit fault prediction DL model without transfer learning, i.e., without parameters initialization in Figure 8, under the same training settings. The learning rate of training is set at $10^{-6}$, and batch size is 32. The training processes are shown in Figure 18.

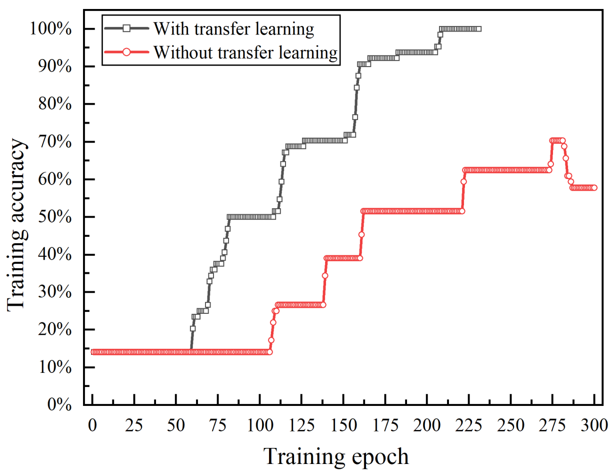


**Figure 18.** Model training process with and without transfer learning.

The heterogeneous transfer learning-based model training approach can make track circuit fault prediction DL model reached 100% accuracy at approximately the 200th epoch, as shown in Figure 18. In contrast, the non-transfer learning-based model training approach under the same conditions only makes DL model with same structure achieves maximum 70% accuracy at around the 270th epoch and then decreases apparently thereafter. That is due to the limited amount of training data available and the presence of noise in the data. This result shows that the transfer learning-based training method can significantly mitigate the issues caused by the inadequate amount of training data and data noise, that can make model training process faster and smoother. Therefore, it can be inferred that the knowledge gained from predicting the RUL of aircraft engine can significantly assist the DL model in the task of track circuit fault prediction.

This paper tested the proposed model using the TC-A test dataset after model training. Figure 19 shows the confusion matrix for the fault prediction result of the dataset.

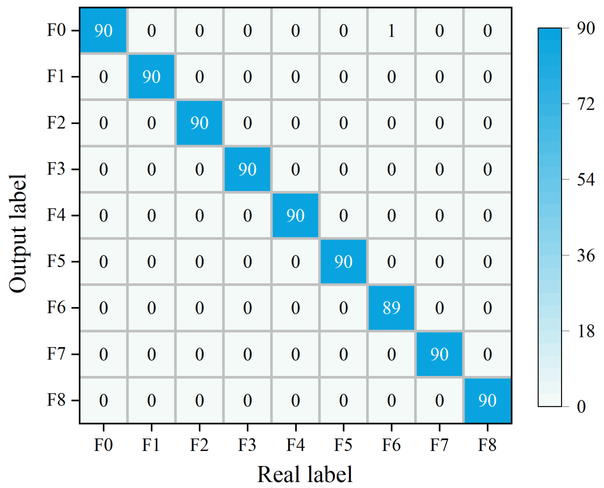


**Figure 19.** Confusion matrix of fault prediction results on TC-A test dataset.

To perform a more comprehensive quantitative analysis of the model's performance, this paper has further calculated four model performance evaluation indices using the provided confusion matrix including accuracy, precision, recall, and F1 score for the DL model of track circuit fault prediction in the TC-A test dataset. The resulting data is presented in Table 4.

**Table 4.** Model performance evaluation indices.

| Index name | Value |
|---|---|
| Accuracy | 99.8765% |
| Precision | 99.8779% |
| Recall | 99.8765% |
| F1 score | 99.8765% |

The model performance evaluation indices results indicate that the DL model proposed in this paper has good performance in all track circuit fault prediction classes, and the prediction result accuracy, precision, recall, and F1 score are all higher than 99%. Thus, it can be concluded that the proposed DL model can satisfy the reliability requirement of the track circuit fault prediction.

To further validate the performance of the proposed track circuit fault prediction model on TC-A test dataset, this paper using the above evaluation indices to compare the proposed track circuit fault prediction DL model with other classical model frameworks, i.e., the Bi-GRU-AM model[28], the Bi-LSTM-AM model[8], the Bi-GRU model[7], the Convolutional Neural Network (CNN) model[33], and the LSTM model[4]. Table 5 displays the performance evaluation indices of different models.

**Table 5.** Comparison of model performance evaluation indices.

| Model name | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| Bi-GRU-AM | 97.77% | 97.77% | 97.88% | 97.78% |
| Bi-LSTM-AM | 95.56% | 95.71% | 95.69% | 95.54% |
| Bi-GRU | 97.78% | 97.88% | 97.78% | 97.79% |
| LSTM | 96.44% | 96.71% | 96.44% | 96.44% |
| CNN | 94.22% | 95.10% | 94.22% | 94.21% |
| **Our model** | **99.87%** | **99.87%** | **99.87%** | **99.87%** |

From Table 5, it is evident that the DL model advanced in this paper surpasses other traditional models in predicting track circuit faults. In addition, the data required for training this model is reduced by 90% relative to training traditional models by transferring expertise from the domain of aircraft engine RUL prediction.

As the TC-A track circuit fault prediction dataset used in this paper is derived from the data collected in the field, and the dataset involves more comprehensive and specific types of track circuit faults as shown in Table 1, it can be assumed that the track circuit fault prediction model designed and implemented in this paper is able to predict most of the common faults that may occur in the future of track circuit in the real operating environment in a real-time high credibility prediction. By applying the relevant methods, it can be realised that the specific location of future faults can be determined before faults occur in track circuit, so as to realise the preventive maintenance of track circuit equipment. So the proposed DL model can further improve the maintenance efficiency and reduces the cost of maintenance of track circuit, while increasing the continuous availability of the relevant equipment.

It is also worth noting that although the training process utilising heterogeneous transfer learning may be more intricate and time-consuming, it can ease the data necessities significantly in the domain of track circuit fault prediction DL model training by utilising openly available datasets from other related research areas. So, the DL model training method proposed in this paper can ability to significantly reduce the workload of actual data collection, so that can greatly enhance the possibility of DL-based IM methods for practical deployment in railway sites so as to enhance the application value of DL algorithms in the field of IM of track circuit practice.

## *Transfer learning effects analysis*

In the preceding subsection, it was initially demonstrated that the DL model training progress in track circuit fault prediction task can be enhanced by utilizing the knowledge acquired from aircraft engine RUL prediction task. In this subsection, we will continue to examine the effect of transfer learning on the training of DL model for track circuit fault prediction. And a more detailed analysis of this impact will be provided by ablation experiments.

Initially, the processes of data distribution fitting network parameters initialisation (DDFN init) and other main layers parameters initialisation (OML init) in heterogeneous transfer learning-based DL model training method shown in Figure 8 is masked, respectively. And then the DL model is trained in the above two cases with all other factors keep constant. The process of training the model is illustrated in Figure 20 for both cases mentioned above.

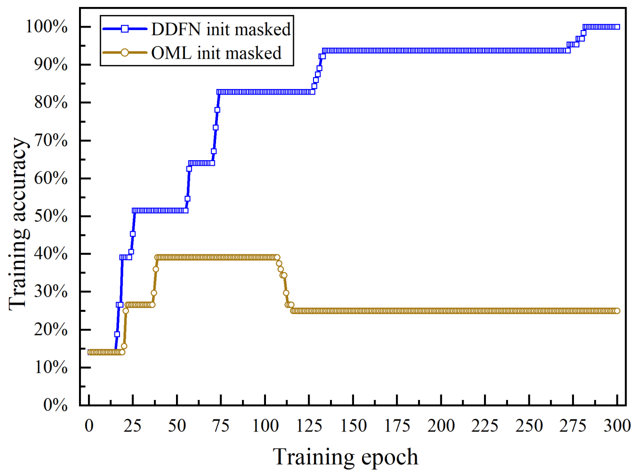


**Figure 20.** Model training process of DDFN init masked and OML init masked.

The model training process with DDFN init masked can make track circuit fault prediction DL model reached 100% accuracy at approximately the 280th epoch, as shown in Figure 20. It can be found that the model delays reaching 100% training accuracy by almost 100 training epochs in the DDFN init masked case compared with relevant results in Figure 18. Therefore, it can be concluded that the DDFN init plays a significant role in accelerating the convergence of model training, and OML init plays a decisive role in the convergence of the model.

It is noteworthy that this model training process does not even perform as well as in the model training without transfer learning shown in Figure 18. Therefore, it could be contended that the isolated process of fitting data distribution creates challenges in the convergence of model training. In this experiment, the process of transforming the data distribution converts the simple feature space in which the data originally exists into a more intricate feature space. As a result, it enhances the nonlinear correlation between the data and makes it increasingly challenging for the DL model training to converge. Consequently, the training performance of DL model deteriorates instead when applying solely the DDFN init.

To further examine the effect of implementing transfer learning to the primary parts of the DL model, we have masked the initialisation process of the parameters of the first Bi-LSTM layer (Bi-LSTM-1 init), the second Bi-LSTM layer (Bi-LSTM-2 init), and the multi-head attention mechanism layer (MAM init) in the DL model, respectively. And then the track circuit fault prediction DL model is trained in these above cases with all other factors keep constant. The process of training the model in all of the above cases is depicted in Figure 21.

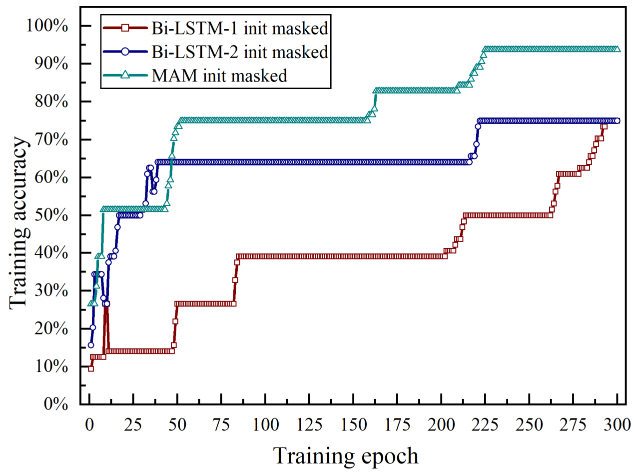


**Figure 21.** Model training process of Bi-LSTM-1 init masked, Bi-LSTM-2 init masked and MAM init masked.

According to Figure 21, it can be found that the early training stages of DL model with MAM init masked and Bi-LSTM-2 init masked outperform the performance even than transfer learning-based training process that have faster and smoother accuracy increase. However, the growth rate of model's accuracy on the training data starts to slow down and reaches a rather low accuracy at last during in the middle to late stages of model training. This implies that although randomly initialized network parameters following a normal distribution can enhance the model's performance in the early stages of training by adapting better to input data, the DL model's final performance can benefit much more from the acquired knowledge in aircraft engine RUL prediction.

It can also be noticed that with MAM init masked, the DL model training speed and final accuracy is rather low compared with transfer learning-based training process. So it can be concluded that the MAM init provides an adjunct to accelerate model training.

## Conclusions and future work

Regarding the important impact of track circuit on the safety and efficiency of railway transportation systems and facing problems that DL based fault prediction models lack sufficient real data for training, the study is carried out on the training methodology of track circuit fault prediction DL models based on heterogeneous transfer learning. This study attempts to propose a transfer learning method that can accelerate the training process of track circuit fault prediction models and reduce the dependence on track circuit data for model training by utilising data from similar domains. Besides, this study also proposes a autoencoder-based index for assessing the feasibility of implementing transfer learning.

The simulation experimental results demonstrate that data from track circuit fault prediction has some similarity with data from aircraft engine RUL prediction in the view of data characteristic. And the transfer learning method proposed in this research can accelerate the model training process with substantially reducing the amount of data used for model training by using aircraft engine RUL prediction data. The relevant researches in this paper can greatly increase the efficiency in the maintenance of track circuit and enhance the relevance of DL-based track circuit maintenance methods.

Future extensions can investigate the generality of the methodology proposed in this study. Also, it is worth studying the methods for transferability assessment.

### Acknowledgements

The authors would like to thank China Railway Shenyang Group Co., Ltd., RTCC of Beijing Jiaotong University and Shuohuang Railway Development Co., Ltd. for offering track circuit microcomputer monitoring data.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

This work was supported in part by the National Key Research and Development Program of China under Grants 2022YFB4300501, the National Natural Science Foundation of China under Grants U2268206, and the National Natural Science Foundation of China under Grants 52102472.

### ORCID iD

Baigen Cai https://orcid.org/0000-0002-6440-005X

### References

1. Yu XY, Dong Y and Dong Y. Fault diagnosis of track circuit based on multi method evidence fusion. *J China Railway Soc* 2021; 43: 86-94.
2. Dai JY and Liu X. Machine learning based prediction of rail transit signal failure: A case study in the United States. *Proc IMechE, Part F: J Rail and Rapid Transit* 2023; 237: 680-689.
3. Hu LQ, He CF, Cai ZQ, et al. Track circuit fault prediction method based on grey theory and expert system. *J Vis Commun Image R* 2019; 58: 37-45.
4. Bruin De T, Verbert K and Babuška R. Railway track circuit fault diagnosis using recurrent neural networks. *IEEE Trans Neural Netw Learn Syst* 2017; 28: 523-533.
5. Xie XX and Dai SH. Fault Diagnosis of jointless track circuit based on deep learning. *J China Railway Soc* 2020; 42: 79-85.
6. Kang XY, Zhao LH, Meng JH , et al. LSTM-based forecasting for number of faults of track circuit compensation capacitor. *J China Railway Soc* 2021; 43: 94-99.
7. Shi XJ and Dai SH. Fault orediction of turnout equipment based on double-layer gated recurrent unit neural network. In: *Proceedings of the 24th IEEE international conference on intelligent transportation systems*, Indianapolis, IN, USA, 2021.
8. Dai SH, Na L and Liang XJ. Research on fault prediction of track circuit based on bidirectional LSTM and attention mechanism. *J China Railway Soc* 2023; 45: 94-102.
9. Fang JS, Chen R, Hu CY, et al. Simulation analysis of high-speed turnout point rail switching force considering influence of external locking device. *Proc IMechE, Part F: J Rail and Rapid Transit* 2022; 236: 48-57.
10. Zhao LH, Guo YM and Klein BD. Analysis of structure importance of compensation capacitor in jointless track circuit. *Proc IMechE, Part F: J Rail and Rapid Transit* 2017; 231: 329-344.
11. Zhu YC, Zhuang FZ, Wang JD, et al. Deep subdomain adaptation network for image classification. *IEEE Trans Neural Netw Learn Syst* 2021; 32: 1713-1722.
12. Pan SJ and Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010; 22: 1345-1359.
13. Tan CQ, Sun FC, Kong T, et al. A survey on deep transfer learning. In: *Proceedings of the 27th international conference on artificial neural networks and machine learning*, Rhodes, Greece, 2018.
14. Feng W, Du J, Zhao Y, et al. A deep learning based data fusion method for degradation modeling and prognostics. *IEEE Trans Reliab* 2020; 70: 775-789.
15. Tang GT, Yi C, Liu L, et al. A novel transfer learning network with adaptive input length selection and lightweight structure for bearing fault diagnosis. *Eng Appl Artif Intell* 2023; 123: 1-10.

16. Su H, Yang X, Xiang L, et al. A novel method based on deep transfer unsupervised learning network for bearing fault diagnosis under variable working condition of unequal quantity. *Knowl Based Syst* 2022; 242: 1-13.
17. Wei W, Zhao XQ and Wu J. Time series prediction of fault rate of high-speed railway on-board equipment based on VMD-ICSO-GRU. *J China Railway Soc* 2023; 45: 58-68.
18. Liu YZ, Zou YS, Wu Y, et al. A novel abnormal detection method for bearing temperature based on spatiotemporal fusion. *Proc IMechE, Part F: J Rail and Rapid Transit* 2022; 236: 317-333.
19. Hochreiter S and Schmidhuber J. Long short-term memory. *Neural Comput* 1997; 9: 1735-1780.
20. Schuster M and Paliwal KK. Bidirectional recurrent neural networks. *IEEE Trans Signal Process* 1997; 45: 2673-2681.
21. Niu ZY, Zhong GQ and Yu H. A review on the attention mechanism of deep learning. *Neurocomputing* 2021; 452: 48-62.
22. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: *Proceedings of the 31st IEEE international conference on neural information processing systems*, Long Beach, CA, USA, 2017.
23. Day O and Khoshgoftaar TM. A survey on heterogeneous transfer learning. *J Big Data* 2017; 4: 1-42.
24. Magotra A and Kim J. Improvement of Heterogeneous transfer learning efficiency by using hebbian learning principle. *Appl Sci* 2020; 10: 5631-5654.
25. Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders. In: *Proceedings of the 25th international conference on machine learning*, Helsinki, Finland, 2008.
26. Borgwardt KM, Gretton A, Rasch MJ, et al. Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics*, 2006; 22: e49-e57.
27. Saxena A, Goebel K, Simon D, et al. Damage propagation modeling for aircraft engine run-to-failure simulation. In: *Proceedings of the 1st IEEE international conference on prognostics and health management*, Denver, USA, 2008.
28. Na L and Dai SH. Fault prediction of track circuit based on bidirectional gated recurrent unit with attention mechanism. In: *Proceedings of the 25th IEEE international conference on intelligent transportation systems*, Macao, China, 2022.
29. Liu KB, Gebraeel NZ and Shi JJ. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. *IEEE Trans Autom Sci Eng* 2013; 10: 652-664.
30. Liu KB and Huang S. Integration of data fusion methodology and degradation modeling process to improve prognostics. *IEEE Trans Autom Sci Eng* 2016; 13: 344-354.
31. Liu KB, Chehade A and Song CY. Optimize the signal quality of the composite health index via data fusion for degradation modeling and prognostic analysis. *IEEE Trans Autom Sci Eng* 2017; 14: 1504-1514.
32. Song CY, Liu KB and Zhang X. Integration of data-level fusion model and kernel methods for degradation modeling and prognostic analysis. *IEEE Trans Reliab* 2018; 67: 640-650.
33. Längkvist M, Karlsson L and Loutfi A. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognit Lett* 2014; 42: 11-24.