Unsupervised anomaly detection system for railway turnout based on GAN

Lei Xue1* and Shuli Gao1

1 School of Computer Science and Engineering, Northeastern University, QinHuangdao 066004, China;
*Corresponding author’s e-mail: xueleineuq@gmail.com (L.X)

Abstract: With the rapid development of society, the railway system plays an important role in human life, and the safety of railways has become an extremely important task. As we all know, the switch is one of the important equipment to ensure the safe operation of trains. Real-time detection of the turnout current plays a vital role in train safety. However, the previous signal-based processing methods require a large number of feature engineering, which greatly increases the workload of pre-processing. Some neural network-based methods show good performance, but for the time series data of the switch, these methods cannot fully extract its local features, resulting in poor information loss and poor prediction accuracy. Based on the generational confrontation network, this paper proposes a kind of unsupervised anomaly detection system. We combine the one-dimensional convolution with the Generative Adversarial Networks (GAN). The one-dimensional convolution network can effectively extract the local features of the time series. The GAN can self-game learning to sample distribution and is better than self-encoder and other models, which improves the accuracy of prediction. In the real railway system, abnormal data is extremely rare and varied, while unsupervised learning does not require label data, and it can well learn the distribution of normal samples. The system improves the efficiency of the staff, accurately diagnoses the switch, greatly shortens the processing time, and avoids the blindness in maintenance. Using our model on the turnout data, the accuracy rate is 0.992, the recall rate is 0.815, and the F1 score is 0.895.

1. Introduction
Under the background of high-speed railway operation, the safety of railway transportation has become the focus of attention [1]. The switch is the weakest part of the track structure and is one of the main factors affecting the safety of railway operation. With the rapid development of railway transportation, the inspection of the switch system has received increasing attention [3]. As a conversion equipment for railway line changes, turnout is the most complicated technical index of the work route and the most demanding line condition. Its reliability has become one of the important factors affecting the safety of railway operation. The detection of the main performance parameters of the turnout is the key means to ensure the safety of the switch and the maintenance. Therefore, the detection technology of turnout has become an important guarantee for the safe operation of trains [2].

By analyzing a large number of turnout curve data, it is possible to judge the electrical and mechanical characteristics of the turnout, and to detect hidden dangers in time. Analysis shows that there may be the following problems: 1) The unlocking current is large, it may be that the locking arc is short of oil, there is jamming when unlocking, and the pressure is high. 2) The operating current is large, which may be a large conversion resistance, such as dirty sliding board. 3) The action is small or
the current is unstable, and the friction band may be loose or fixed, and the node in the starting circuit is in poor contact. 4) The locking current is large, which may be tightly attached, sharp rails and foreign objects.

At present, there are many methods for detecting railway turnouts. In order to realize the fault diagnosis of turnouts, scholars at country and abroad have proposed many different methods and achieved important results. The data integration module based on cubic spline interpolation is used to classify the turnouts action. The new RBF neural network is used for fault diagnosis [4]. In the [14], based on the collected instantaneous current values, a fault diagnosis method for railway turnout based on support vector machine is proposed. In [5], wavelet analysis is applied to the analysis of the action current curve of the motor switch machine to determine the fault types corresponding to various feature vectors. Due to the redundancy of features in the process of feature extraction, [6] proposed to reconstruct the current curve by discrete wavelet transform to obtain the information of each frequency band, and then use the principal component analysis method to select the effective frequency band information feature vector to reduce the feature dimension, then BP neural network is used for current classification. However, all of the above methods require manual extraction of a large number of features. This may be imperfect in a real-time monitoring system, and this often loses some important information, thereby reducing the accuracy of prediction.

With the rapid development of neural networks, some new detection methods have emerged. The emergence of neural networks simplifies the work of many feature extractions and has strong learning and generalization capabilities. [7] Proposed an improved BP algorithm, set up a parallel neural network for information fusion, and set a double threshold method for fault diagnosis. [8] Proposed a BP neural network algorithm, by summarizing the typical speed-increasing fault current curve and extract the characteristic vector value of the action current curve for fault diagnosis. [9] The LSTM network is used to automatically extract features, and then the network classifier is used to classify faults according to features. Although the above methods make full use of the advantages of the automatic extraction feature of the neural network, none of these methods can well collect the local information in the time series of the switch, and the abnormality usually appears in the local information. The above methods all use supervised learning methods. However, in our actual scenario, the abnormal situation is uncontrollable, and the number of abnormal samples is extremely rare. Given this situation, it is very reasonable and effective to use an unsupervised method for anomaly detection.

The turnout current data we studied belong to the one-dimensional time series category, and most of the previous anomaly detection methods based on the Generative Adversarial Networks are applied to the field of image recognition, which is not well applied to the problem of turnout detection. Recently, Shaojie Bai et al. [13] applied one-dimensional convolution to sequence problems and achieved more significant results than networks such as LSTM. We are inspired by this, drawing on the work of others, and applying the one-dimensional convolution network to the Generative Adversarial Networks for the problem of the time series of the switch. The architecture of the network can learn more subtle features than the model such as the autoencoders. And the one-dimensional convolution can fully extract the local features of the time series without losing its time dependence, and use the one-dimensional convolution generation against the network to detect the switch current curve. Our model contains two indicators: first, the residual error is calculated by the generator looking for the reconstruction error of the potential sample space and the real sample space. Second, the discriminator's score is used to calculate the discriminant loss. Our model is based on the weighting of the two indicators to detect abnormality.

2. Model description
The first proposal of the Generative Adversarial Networks is applied in the image field, and the training process is similar to the two-player game to train the two networks [10]. Our goal is to train an unsupervised anomaly detection network. Previously [15] used the basic framework of [17] to perform anomaly detection in the image field and achieved remarkable results.
The railway turnout data belongs to the one-dimensional time series. We draw on the ideas of the predecessors [15], using the combination of one-dimensional convolution and GAN, so that the original features of the data can be highly maintained and effectively extracted to its important parts feature. The network structure is shown in Figure 1. The generator $G$ is a standard one-dimensional convolution network in which the generator $G$ maps the potential spatial variable $z$ through the generator to the original data space by learning the distribution $G(z)$ of the data $X$ on the training set $M$ (normal sample), using a vector $z$ with a dimension of $1 \times 10$. The first four layers of the network use a step size of 2, the convolution kernel size is set to $1 \times 5$, the number of convolution kernels is 20, 40, 80, and 160, respectively. The last layer uses a $1 \times 1$ convolution kernel. The activation function use relu. The discriminator $D$ also use one-dimensional convolutional network, and the output $D(\cdot)$ of the discriminator can be interpreted as the true time series information of training data $X$ sampling or the true and false probability $G(z)$ generated by the generator $G$, where the convolution kernel size is set as $1 \times 5$, step size is 2, sigmoid activation function is used in the last layer, and Leaky_relu is used in the remaining layers. $G$ and $D$ optimize through the game of both parties with the most valuable function below [10].

$$\min_{G} \max_{D} V(D, G) = E_{x \sim p_{data}} \left[ \log D(x) \right] + E_{z \sim p_{z}} \left[ \log \left(1 - D(G(z))\right) \right]$$

(1)

Next, we describe the loss of the model, which consists of two parts [17]: residual loss and discriminator loss. Based on generation $G(z)$, we define a loss function to map new data to potential space. Residual loss is to measure the numerical difference between the original data $Q$ and the generated data $G(z)$ in the generated sample space. It is defined as:

$$L_{R} = \sum |Q - G(z)|$$

(2)

Given the real sample $Q$, our goal is to find points $z$ corresponding to the feature information $G(z)$ in the potential space, which is closest to the real sample $Q$ numerically. In order to find the optimal point $z$, we randomly sample $z$ from the potential spatial distribution $z$ and feed it to the training generator to obtain the generated signal $G(z)$. The residual loss enhances the numerical similarity between the generated data $G(z)$ and the real data $Q$. The value of $z$ is adjusted by using the modules after training, discriminator $D$ and generator $G$, and the coefficient of $z$ is self-adapted by back-propagation. The training parameters of generator and discriminator remain unchanged.

Because we use unlabeled data when we train networks, our goal is not to learn class-specific identification characteristics, but to learn good representations of them. Yeh et al. proposed a feature matching based identification loss [16], in which the objective function of the optimized generator was used to improve GAN training. Therefore, in antagonistic training, we do not adjust the training target of the generator, but use the idea of feature matching to improve the mapping of potential space.
Instead of using the scalar output of the discriminator to calculate the discriminator loss, we use the richer intermediate feature representation of the discriminator and define the discriminator loss as follows:

$$L_D(z) = \sum |f(x) - f(G(z))|$$ \hspace{1cm} (3)

Here we use the third layer of the discriminator to calculate the discriminator loss. We defined the total loss as the weighted sum of two components, which was set $\lambda$ as 0.3 in this experiment:

$$L(z) = (1 - \lambda) L_D(z) + \lambda L_P(z)$$ \hspace{1cm} (4)

When anomaly detection is carried out on new data, formula (4) is used to find $z$ that maps new data to potential space, and we also use the score of this formula to identify exceptions.

3. Experiments and results

Under normal circumstances, there are few faults in the railway switch, so abnormal data is very rare for us, but normal data can be easily obtained. In this paper we collect 904 data, 17 of which are abnormal data. In order to ensure the diversity of samples and the effectiveness of the algorithm, we randomly divide the training set and test set from the sample data. Since the input of the neural network requires the unification of the data dimensions, the real data dimension of the switch can be distributed between 100-200 dimensions, so we simply fill and intercept all the data to make it have a uniform dimension. Because we use one-dimensional convolution, it can collect the features of each part without losing information, so it is not necessary to perform complex preprocessing such as [4].

We also compared the previous SVM [14], LSTM [9], and replaced the one-dimensional convolutional network in the model with a common fully connected layer network (F-GAN) to compare with our model. This article builds an anomaly detection model using Tensorflow2.0 based on the Python 3.6 environment. The learning rate is set to 1e-4 and optimizer use the Adam. The final experimental results are shown in Table 1.

| Method   | Pre  | Rec  | F1   |
|----------|------|------|------|
| SVM      | 0.631| 0.562| 0.595|
| LSTM     | 0.916| 0.634| 0.749|
| F-GAN    | 0.921| 0.624| 0.744|
| Our Model| **0.992**| **0.815**| **0.895**|

Analysis of Table 1 shows that the performance of the LSTM architecture on such issues is significantly better than the SVM framework, which also underscores the superiority of LSTM to time series problems. The F-GAN framework is slightly better than the LSTM. But there is no obvious optimization. Our model has a significant improvement over the other three, which also shows that the one-dimensional convolutional generation of GAN has excellent performance in the problem of turnout.

4. Conclusion

A large amount of turnout data can be collected in the current railway physical system, which can be used to monitor the behavior of the system to detect various anomalies of the railway switch. The results show that the combination of one-dimensional convolution and the GAN shows performance superior to other baseline methods in the detection of turnouts. For the detection of turnout faults, this real-time detection system can find the fault of the turnout in the first time, and has important practical and social significance in ensuring the safety of the railway.

Acknowledgement

The author would like to thank the Northeastern University Qinhuangdao Branch for the support of the machine learning and gaming laboratory, and at the same time thank the relevant personnel for their fruitful and profound discussion and guidance.
References

[1] Wahid F, Allan M, Gerard V E, Thiru A, Sakdirat K, Alex R. (2015) Composite railway sleepers – Recent developments, challenges and future prospects[J]. Composite Structures, (134):158-168.

[2] YuNing Zhang, Qi Xie. (2015) Railway vision detection method based on machine vision [J]. Computer applications and software, (1):225-228.

[3] Cong Gao, WenXiang Zhou, YuanBin Zhang. (2016) Current Status and Development of Railway Turnout Detection Technology [J]. mechanical, 43(7):34-38.

[4] Huan Mei, YanDong Ma, JiuSi Shan, et al. (2017) Switch fault diagnosis system based on spline interpolation and RBF network [J]. Computer technology and development, (5).

[5] Yu Dong, YuanYuan Zhao, HaiXiang Lin. (2012) Analysis and Research on the Action Current of Electric Switch Machine Based on Wavelet Analysis [J]. Journal of Lanzhou Jiaotong University, 31(6):39-43.

[6] Xu Zhang, WenJun Wei. (2018) Identification of Turnout Fault Current Characteristics Based on PCA-WT [J]. Journal of Lanzhou Jiaotong University, v.37: No.187(2):49-54.

[7] TieJunWang, Yu Dong, ChengLu Shen, et al. (2010) Research on Intelligent Fault Diagnosis Method of Turnout Based on Neural Network [J]. West Rail Technology, (2):8-10.

[8] YongFeng Ju. (2014) Fault Diagnosis Algorithm for Speed Increased Turnout Based on BP Neural Network [J]. Journal of Wuhan University of Technology (11):77-81.

[9] WeiHuaTang. (2019) Research on Extraction and Recognition of Turnout Fault Features Based on LSTM/NN [J]. Computer applications and software, 36(01):165-169.

[10] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. (2014) Generative Adversarial Networks [J]. Advances in Neural Information Processing Systems, 3:2672-2680.

[11] Schlegl T., Seeböck P., Waldstein, S. M., Schmidt-Erfurth, U., & Langs, G. (2017, June). Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In International Conference on Information Processing in Medical Imaging (pp. 146-157). Springer, Cham.

[12] LiZhe Yao. (2014) Principle and Analysis of ZD6 Ballast Current Curve Monitoring [J]. Enterprise technology development: Mid-term publication, (5):79-79.

[13] Bai S, Kolter J Z, Koltun V. (2018) An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[J].

[14] Xing Zhang, WenJun Wei. (2016) Switching fault action current identification based on support vector machine [J]. Computer knowledge and technology, 12(7):264-265.

[15] Schlegl T, Seeböck P, Waldstein S M, et al. (2017) Unsupervised anomaly detection with generative adversarial networks to guide marker discovery[C]//International Conference on Information Processing in Medical Imaging. Springer, Cham: 146-157.

[16] Yeh, R., Chen, C., Lim, T.Y., Hasegawa-Johnson, M., Do, M.N.: Semantic image inpainting with perceptual and contextual losses. arXiv:1607.07539 (2016)

[17] Radford A, Metz L, Chintala S. (2015) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J]. Computer Science.