https://doi.org/10.31449/inf.v49i19.7872 Informatica 49 (2025) 89–96 89

Deep Learning-Based Defect Identification for Transmission Tower Bolts:Optimization of YOLOv3 and ResNet50 Algorithms

Huiwei Liu[1], Pengjie He[1], Ziqiang Lu[2*], Jie Li[1], Ziying Lu[1]
[1]STATE GRID UHV TRANSMISSION CO. OF SEPC, Taiyuan, Shanxi 030001, China
[2]No. 12, Nanxiaoqiang, Xinghualing District, Taiyuan, Shanxi 030001, China
[*]E-mail: lzq_qiang@hotmail.com
[*]Corresponding author

Keywords: transmission tower, bolts, defect identification, deep learning

Received: December 21, 2024

In the power industry, it is often necessary to carry out inspection on the transmission tower, including the detection of bolts on the tower. In the past, the method of detecting whether the bolts are tightened is manually completed, which has a large workload and low efficiency. Therefore, this paper proposes to use deep learning method to detect and identify the defects of bolts on the tower to improve the detection efficiency. In bolt detection, the data enhancement method is used to increase the number of samples. In the link of bolt detection, YOLOv3 algorithm is used. In order to improve the detection accuracy of the algorithm, the algorithm is also optimized. In the identification of bolt defects, ResNet50 network, data enhancement and transfer learning are adopted to solve the problem of bolt defect identification, and the ResNet50 network is optimized to improve the recognition quality of the algorithm. The recognition accuracy of nut and screw defects is 0.95 and 0.90 respectively. The feasibility of the identification method is confirmed, which can be used to identify bolt defects of transmission tower and improve the detection efficiency and quality of transmission tower.

Povzetek: Članek predstavlja metodo globokega učenja za identifikacijo nepravilnosti na vijakih na prenosnih stolpih, npr. premalo zategenjeni. Optimizirana sta algoritma YOLOv3 in ResNet50 za zaznavanje napak na vijakih, kar povečuje točnost detekcije do 95 % za matice in 90 % za vijake ter izboljša učinkovitost in kakovost preverjanja prenosnih stolpov.

1 Introduction

In the era of rapid development, electricity consumption is increasing day by day, and more and more transmission lines need to be laid by the power industry. Therefore, the expectation of the state and people for the further construction of the power grid is also increasing, and the power industry is also constantly seeking development [1]. To ensure the normal use of electricity, China continues to improve the transmission level of high voltage and UHV on the basis of the original, and the total length of the laid transmission lines continues to increase, so the number of transmission equipment that needs to be operated continues to increase [2-4]. Transmission poles and towers play a key role in the entire transmission system, which supports the effective circulation of transmission lines, so the carrying capacity of transmission poles and towers directly determines whether the entire transmission line can operate safely and stably [5-6]. For the existing transmission tower, in order to facilitate the construction personnel in the complex and changeable site for efficient assembly, it often utilizes an angle steel connection structure. Therefore, when the tower is in a strong wind environment above level 5 for a long time (6 months to 1 year), the bolts connecting each node will slowly loosen [7]. If it is not found and treated in time, and the number of bolts loosening and falling off exceeds 1/3, then the tower will tilt and collapse, resulting in an electrical accident. [8]. In order to reduce the occurrence of such problems, this paper will explore the defect identification method of transmission tower bolts, aiming to find out the current and efficient bolt defect identification method, reduce the occurrence of electricity accidents caused by bolt loosening, and protect the transmission industry.

2 Theoretical overview

2.1 Deep learning methods for computer vision

Deep learning technology has achieved considerable application results and value in many fields. As the basis of deep learning, neural network is an algorithm model that is regarded as imitating biological neural network, so it is also called artificial neural network [9-10]. Similar to biological neurons, neural networks are made up of many neurons that contain a certain number of inputs and outputs. When many similar neurons are connected to each other, a basic neural network is formed, its structure consists of three parts: one is the input layer; The second is the output layer; The third is the hidden layer. If the activation function is not introduced into the neural network, it is only a linear combination for the neural network, no matter how many layers there are. Because of the existence of the activation function, the forward propagation can represent a nonlinear combination, so that the neural network has the ability of nonlinear expression. Among them, the Sigmoid function, as the earliest activation function, has an output range of 0 to 1, and its derivative value is always less than 1. Therefore, when the network is relatively deep, its partial derivative value will continue to decrease, making the network gradient disappear. The derivative of the ReLu function is 1 when x is positive, and 0 when it is not, thus avoiding the gradient disappearance of the deep network. Therefore, when building a relatively deep network, the ReLu activation function is commonly used. When the network is transmitted forward, the desired output can be obtained by intervening with the neural network., which is often deviated from the real value. In order to make the result of forward propagation reasoning as close as possible to the real value, under the backpropagation algorithm, the parameters in the network will be updated and optimized, the network performance is effectively enhanced. In the neural network, the error of the output result of the network can be understood by observing the change law of the loss function. At the same time, under the function, the network parameters will be updated and optimized. After several backpropagation, the optimization weight of network parameters is higher, and then the quality of network prediction is improved, and the prediction accuracy is improved. It is this characteristic that can better cope with the complex environment of the transmission tower, and can capture the state of the tower bolt and escort the tower for operation.

2.2 Object detection technology based on deep learning

Object detection technology A deep learning detection technology that detects the position information and category information of the input image, and then transmits the detected results to the terminal. The algorithm can be subdivided into two algorithms according to different stages: One is a two-stage target detection algorithm. The first stage is to obtain the target candidate frame, and the second stage is to conduct classification processing, mainly represented by R-CNN series and SPP-Net (Spatial Pyramid Pooling in Deep Convolutional Networks). The second is a single-stage target detection algorithm, which can directly obtain the location and category information of the target without segmented processing, mainly represented by YOLO and SSD (Single Shot Detector). Among them, YOLO in addition to the detection speed, but also has the advantages of low background false detection rate, more suitable for transmission tower monitoring. Since its release, YOLO has undergone several iterations and improvements, forming a number of versions, such as YOLOv1, YOLOv2, YOLOv3, etc. For YOLOv3, the feature pyramid network (FPN) is added, which improves the detection ability of different scale targets, and can better cope with the actual situation of transmission towers. Therefore, selects YOLOv3 algorithm to solve the detection problem of pole and tower bolts. YOLOv3, the third version of YOLO series, belongs to the single-stage target detection algorithm. Compared with the previous generation of YOLOv2, YOLOv3 improves the prediction accuracy while maintaining the speed advantage, especially strengthening the recognition ability of small objects. YOLOv3 has made improvements in three aspects: First, the number of convolutional network layers is increased; Second, the grid structure of FPN is used for reference, and multi-scale feature detection is used to strengthen the detection ability of the algorithm for small and medium-sized targets in images. Third, the cross-entropy loss function is used instead of the mean square error.

3 Grid structure design and training optimization strategy

3.1 Grid structure design

The ResNet50 network uses 1x1 convolution for downsampling, which results in the loss of information in the feature graph and affects the model performance. Therefore, this paper uses the model fine-tuning method to improve the existing ResNet50 network:1) Improved 7x7 convolution: The first convolution layer of ResNet50 uses a 7x7 convolution kernel, which is used to adjust the number of input image channels to 64 and the length and width to 1/2 of the original. In Inception v2, it was proposed that multiple small convolution kernels could be stacked to obtain the receptive field of large convolution kernels, and this approach was also applied in VGGNet. By stacking two 3*3 convolution nuclei, you can get the receptive field of 5*5 convolution nuclei, and this law can be similar to that, for example, when the number of 3*3 convolution nuclei superposition reaches 3, you can get the receptive field of 7*7 convolution nuclei. Using multiple small convolutional stacks instead of large convolutional kernels can reduce the number of parameters in the network and improve the expressibility of the network. Therefore, the superposition law can effectively reduce the number of network parameters and improve the network expression ability. Therefore, this paper will use this way to replace large convolution kernel, such as using three 3*3 convolution kernels instead of 7*7 convolution kernel. 2) Improvement of the subsampled convolutional layer: In the convolutional neural network, the subsampled convolutional layer is one of the conventional methods, and the advantages of this method lie in the small number of parameters, high degree of fit, and low computational cost, as shown in Figure 1.

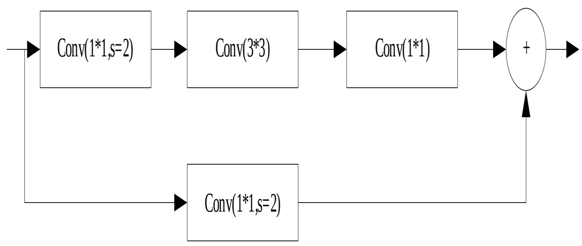


Figure 1: ResNet50 fuses downsampling operations into the residual structure resulting in 3/4 information loss

ResNet50 fuses the downsampling operation into the residual structure, so that although the size of the feature graph is reduced, the 1x1 convolution kernel will cause 3/4 information loss when the downsampling operation with step size 2 is performed, in this lost information, it is easy to mix more details. which is not conducive to the judgment of bolt defects. However, the difference between normal bolts and defective bolts is small, and it is often necessary to judge the category by the details in the image, so the loss of details will directly affect the final judgment of the model. To ensure the comprehensiveness of the feature map information, the residual structure including the downsampling operation is improved in this paper, as shown in Figure 2.

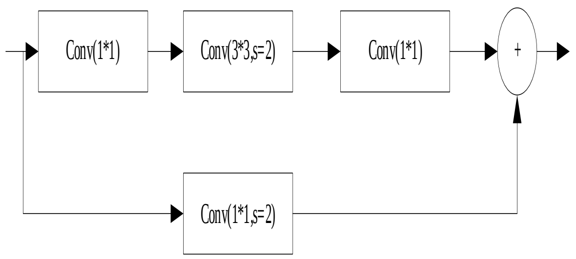


Figure 2: The residual structure with downsampling operation is improved to reduce the information loss in the feature graph

It is not difficult to find that if the step size of 1*1 convolution kernel and 3*3 convolution kernel is 1 and 2 respectively, and the set convolution kernel is transferred to 3*3 convolution kernel by means of downsampling method, it can effectively reduce the loss of detail information in the feature map.

3.2 Training optimization strategy

3.2.1 Transfer learning

The main way to use transfer learning in the field of deep learning is to train convolutional neural networks such as VGGNet, ResNet and GoogLeNet on data sets such as ImageNet to obtain training weights, and then apply the training weights to the target data set to complete the classification task. Due to the large sample size of lmageNet data, it is necessary to train the model to improve its generalization level and cater to the processing of subsequent classification tasks. Compared with the lmageNet data set, the data collection of transmission tower bolts has a large expenditure and a small sample size. Without intervention, overfitting problems are likely to occur during training, which will affect the subsequent output results. In this paper, a transfer learning method is proposed to identify bolt defects. After obtaining the pre-training weight, it is loaded onto the bolt data set, and the parameters are initialized to improve the training effect.

3.2.2 Learning rate adjustment

In the beginning, in order to speed up the training speed and make the parameter update quickly, a relatively large learning rate is often set at the beginning. At this time, the loss decreases rapidly as each parameter update step grows longer. However, using too large learning rate is likely to lead to unstable training, and it is easy to skip local optimal solutions. Therefore, the learning rate needs to be reduced, this is usually done using step delay, which is to multiply the learning rate by the gamma coefficient when the training round reaches the specified round. However, this strategy also has a problem, because the weight parameter is far away from the final solution at the beginning of training, using too large learning rate may lead to the instability of the value obtained by training. Therefore, to avoid this problem, warm up is often used to intervene to change the learning rate state, even if it changes to the initial state, as shown in Formula 1.

$$\mathrm{lr_i} = \mathrm{i}^{*}\eta/\mathrm{m}(1)$$

Where, the initial learning rate is represented by η; When the training cycle i is equal to m, the initial learning rate is reached. Although the way of gradual attenuation reduces the learning rate in the training process, its transition is relatively blunt. Therefore, the learning rate attenuation method of cosine annealing is proposed by taking advantage of the smooth decline of cosine function. As shown in formula 2.

$$\mathrm{lr_i} = \frac{1}{2}\left(1 + \cos\left(\frac{\mathrm{t}\pi}{\mathrm{T}}\right)\right)\eta(2)$$

Where, the training rounds are represented by T, the current training rounds are represented by t, and the maximum learning rate is represented by η. This method takes advantage of the fact that the cosine function changes slowly near the domain 0 and π, but changes faster near $\frac{1}{2}\pi$, so the learning rate of this method changes more diversified and smoother. To sum up, the learning rate adjustment strategy of preheating and cosine annealing was finally adopted to optimize the training process. Let the maximum and minimum learning rates be 0.001 and 0.0001, respectively (The learning rate is not less than 0). First, the learning rate is preheated, the learning rate is increased by 0.0001 each time, and the maximum learning rate is obtained in the 10th round, and then the cosine annealing learning rate is adjusted, T is 90.

3.2.3 Data enhancement

The sample size of the obtained screw and nut dataset is relatively small, and for deep learning, too small sample size can lead to overfitting problems. For overfitting, the simplest and most effective method is to increase the training samples of the input network, so the classification data set needs to be enhanced to get more training samples. The amplification methods of bolt training samples include brightness and contrast transformation, rotation, noise disturbance, flipping and other data enhancement methods. The conversion formula of brightness and contrast as follows:

$$g(x,y) = \alpha f(x,y) + \beta (3)$$

Where, the image pixel values before and after the intervention are represented by $f(x,y)$ and $g(x,y)$ respectively. By adjusting α you can change the image contrast, by adjusting β you can change the image brightness. The so-called image rotation is rotated around a point at a given Angle. Assuming that the coordinate of the pixel before rotation is (x,y), the coordinate after rotation is $(x^{`},y^{`})$, and the rotation Angle is θ, then the formula for image rotation is as follows:

$$x^{`} = \cos\theta x + (-\sin\theta)y(4)$$

$$y^{`} = \sin\theta x + \cos\theta y(5)$$

Random addition of Gaussian noise or salt-and-pepper noise to the image can be used to simulate the noise obtained from images taken in different environments. Among them, Gaussian noise is a random number, which follows the probability distribution Gaussian distribution and satisfies all channel pixels. Salt and pepper noise refers to randomly setting the pixel value of a certain pixel in the image to 0 or 255°. Turns can be divided into vertical flip and horizontal flip. Among them, the symmetry axis of the vertical flip is the X-axis, and the pixels on both sides of the image can be exchanged during the flip. The symmetry axis of the horizontal inversion is the Y-axis, and the pixels on both sides of the image can be exchanged during the inversion.

4 Experimental research

4.1 Transmission tower bolt defect identification process

The transmission tower bolt defect identification process is divided into three parts: first, the unenhanced original data set is used for training on the unimproved ResNet50 network; second, the optimization strategy is used for training on the unimproved ResNet50 network; third, the improved ResNet50 network is used for training, the defect identification process of transmission tower bolts is shown in Figure 3.

In addition, all experiments were trained separately on both the nut and screw datasets, and both were trained under the same conditions. The experimental configuration of the experimental training platform used as follows:

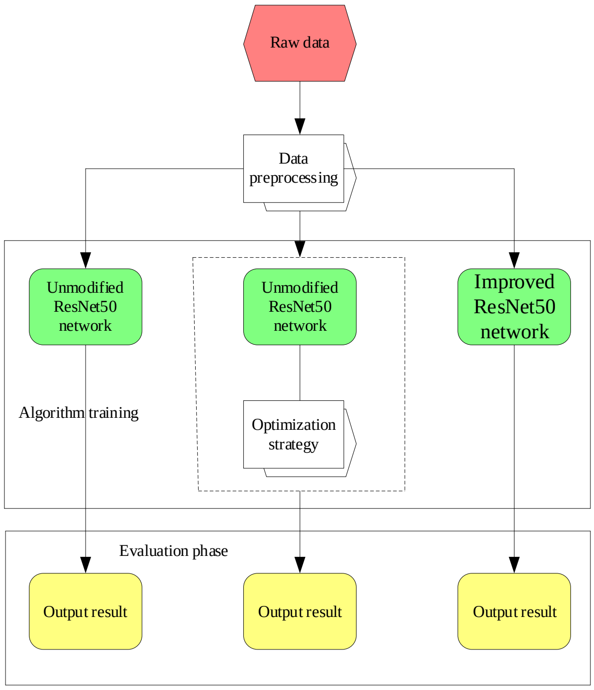


Figure 3: Defect identification process of transmission tower bolts

Table 1: Experimental configuration

Name	Argument
GPU	NVIDIA RTX 3090
CPU	Intel@CoreTMi9-10900k
Operating system	Ubuntu20.10
Internal memory	64G
Development voice	Python3.7
Deep learning framework	Pytorch1.7

4.1.1 Training experiment of the original data set

The unimproved ResNet50 network was used to train the unenhanced nut and screw classification data set, and the ImageNet pre-training weights were not used to train the model and 100 rounds of training were performed, with 64 input images for each iteration training, the loss function used is the loss function of YOLOv3. The Adam optimizer is used when setting the learning rate parameter. During the training period, the loss change results are shown in Figure 4.

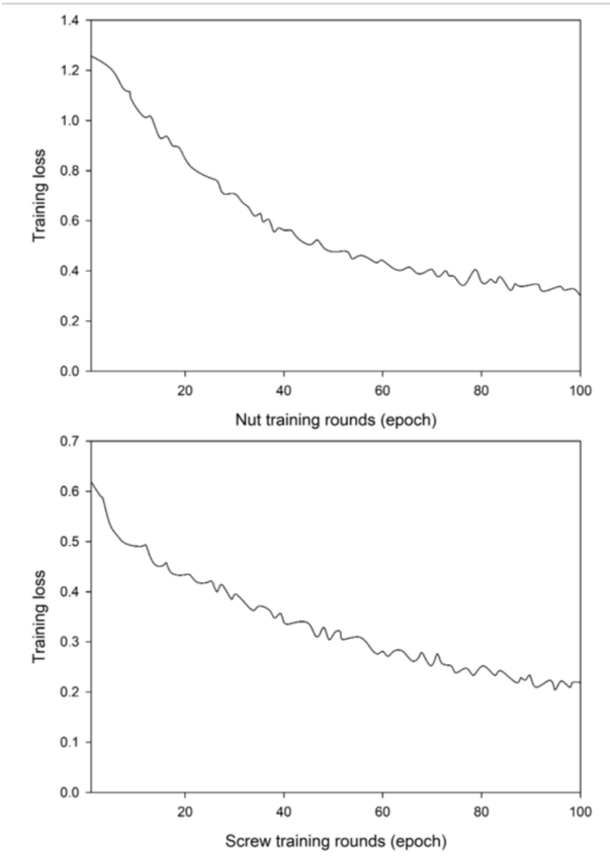
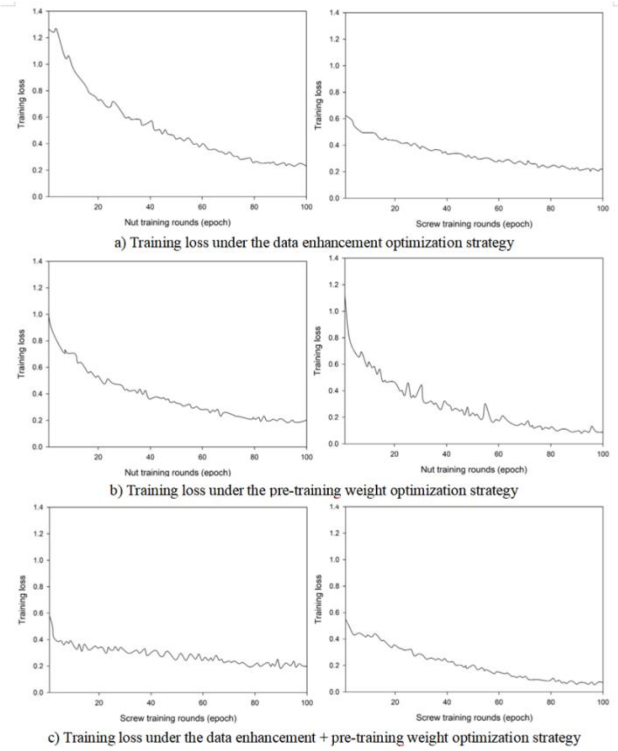


Figure 4: Experimental training loss of basic scheme

In the graph, the horizontal coordinate represents the current training rounds, and the vertical coordinate represents the training losses.

4.1.2 Optimization strategy application experiment

Three comparison experiments were conducted using the unmodified ResNet50 network. In the first comparison experiment, the data sets of nuts and screws after data enhancement were trained separately, without using pre-training weights, the loss function is the same as above, which is still the loss function of YOLOv3. The second comparison experiment was trained on the data set without data enhancement, using the pre-training weight; The third comparison experiment was conducted on the data set after data enhancement, using pre-training weights. In the above experiments, 100 rounds of training are required for the model and 64 input images are required for each iteration training. Using Adam optimizer, the learning rate adjustment mechanism using preheat and cosine annealing, maximum and minimum learning rates are set to 0.001 and 0.0001 (learning rates greater than 0). The initial intervention in the learning rate is made by preheating treatment, and the learning rate is successively increased by 0.0001, and the maximum learning rate is obtained in the 10th round of training, and then the cosine annealing learning rate is adjusted, the period is 90. The loss changes in the training process of each comparison experiment are shown in Figure 5.



a) Training loss under the data enhancement optimization strategy

b) Training loss under the pre-training weight optimization strategy

c) Training loss under the data enhancement + pre-training weight optimization strategy

Figure 5: Training loss of the experiment with optimization strategy

4.1.3 Experiment on improving network structure

The improved ResNet50 network was used to train the nut and screw classification data set after data enhancement, and the ImageNet pre-training weights were used to initialize the model and perform 100 rounds of training, with 64 input images for each iteration training. Using Adam optimizer, the learning rate adjustment mechanism using preheat and cosine annealing, maximum and minimum learning rates are set to 0.001 and 0.0001 (learning rates greater than 0). The initial intervention in the learning rate is made by preheating treatment, and the learning rate is successively increased by 0.0001, and the learning rate reaches the maximum value in the 10th training round, and then the cosine annealing learning rate is adjusted, the period is 90. Figure 6 shows the change of loss during training.

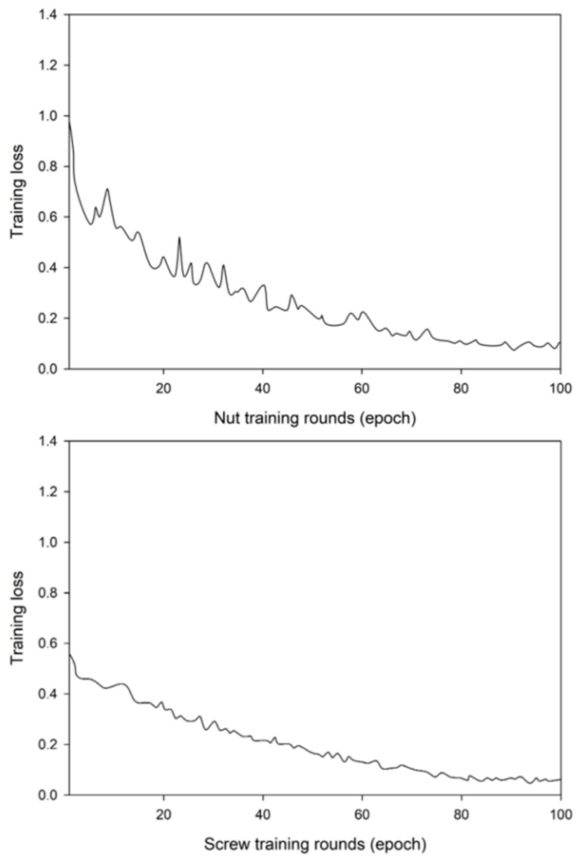


Figure 6: Training loss of the improved network structure experiment

In the graph, the horizontal coordinate represents the current training rounds, and the vertical coordinate represents the training losses. From the training results, it is found that the training loss of both nuts and screws decreases with the increase of training rounds. When the training rounds reach 100, the training instantaneous is between 0-0.2. On the whole, training rounds and training losses are inversely related, increasing and decreasing.

4.2 Result analysis

Use the trained network weights to test on the test platform. The system Windows10 used in the test platform, the processor is Intel i7-8750H, the memory is 16G, and the graphics card is GTX1050 MAX-Q. Random extraction was carried out for each category test set by random extraction, and the number was set to 50. After the extraction, the images were trained and the training weights were obtained for prediction. The results were shown as follows:

Table 2: Identification results

Option	Nut recognition accuracy rate	Correct rate of screw identification
Basic scheme	0.83	0.80
Data enhancement	0.89	0.86
Pre-training weight	0.90	0.84
Data enhancement + pre-training weight	0.92	0.88
Improved network structure	0.95	0.90

In the basic scenario, the original data set without data enhancement was directly used for training on ResNet50. Because few samples were trained, and no learning adjustment and transfer learning interventions were added, the recognition effect of the algorithm was not ideal, and the obtained recognition accuracy rate was 0.83 for nuts and 0.80 for screws. Using data enhancement alone, it was found that the nut recognition accuracy was 0.89 and screw recognition accuracy was 0.86, which were improved by 0.06 and 0.06, respectively, compared with the basic scheme. After the intervention with the pre-training weight, the nut recognition accuracy is 0.90 and screw recognition accuracy is 0.84, which is 0.07 and 0.04 higher than the basic scheme, respectively. Methods such as data enhancement, transfer learning and learning rate adjustment were used for training. As the number of samples participating in training was increased, the training method was optimized, enhanced recognition quality, with the nut recognition accuracy of 0.92 and screw recognition accuracy of 0.88. Using the improved network for training, after network optimization and adjustment, the corresponding model has a strong expression ability, which reduces the probability of feature image information loss to a certain extent. Compared with the network without optimized adjustment, the recognition accuracy will be improved, and the results confirm this view. For example, the recognition rate of nuts and screws is 0.95 and 0.90, respectively, which has higher recognition strength. In summary, using data enhancement, transfer learning, learning rate adjustment and improving the ResNet50 network can comprehensively improve the recognition accuracy of the algorithm.

5 Summary

To sum up, this paper interferes with the data by means of data enhancement technology, transfer learning method and learning rate adjustment, which makes the data sound and stable and establishes a basis for subsequent bolt defectidentification. In addition, the ResNet50 network is improved, which further strengthens the expression ability of the network and guarantees the soundness of the feature map information. From the empirical results, the initial recognition rate of nuts was only 0.83, and after optimization, the recognition rate reached 0.92; The screw was increased from 0.80 to 0.88, which confirmed the feasibility of the scheme. After further improving the network structure, it is found that the recognition rates of nuts and screws reach 0.95 and 0.90 respectively, it means that the method can improve the efficiency and quality of bolt defect identification.In terms of novelty, it breaks the traditional core setting and replaces it with 7x7 core, reduces the number of parameters in the original network, improves the expression ability of the network, and makes a paving for the identification accuracy of the transmission tower.At present, only tower bolts have been identified, and future studies will identify more categories such as tower insulators and bird's nests. At the same time, there is still room for improvement in the recognition accuracy rate. In the future, we will try to combine traditional image processing methods to further improve the recognition performance of the algorithm.

Funding

This study was supported by State Grid Shanxi Electric Power Company Technology Project: Research and Application of an Online Monitoring System for Bolt Looseness in Transmission Towers Based on DSP Voiceprint Information (52051K240001).

References

[1] Zhou D (2024). Overview of the development of China's green power industry. *China and Foreign Energy*, 29(03), pp. 76.

[2] Yang H, Zhou H, Gao B, et al. (2025). Non-contact sensing method for bolt loosening defects of transmission tower. *Yunnan Electric Power Industry*, (01), pp. 20-25+36.

[3] Xie S (2023). Hidden Danger automatic detection Method of High voltage Transmission line based on Deep Learning Algorithm. *Manufacturing and Upgrading Today*, (12), pp. 37-39.

[4] Wang G, Liang Y (2023). Research on the application of new energy-saving wires in high voltage transmission lines. *Energy Conservation and Environmental Protection*, (11), pp. 58-62.

[5] Li B, Li X, Hou S, et al (2021). Surface electric field Calculation Method of 5G Communication Equipment Based on HV transmission line sharing tower. *Southern Power Grid Technology*, 15(10), pp. 65-71. DOI: 10.13648 / j. carolcarrollnki issn1674-0629.2021.10.009.

[6] Wang T, Yan L, Chen D, et al. (2018). Research status of Power transmission line tower Stress analysis. *Science and Technology Innovation and Application*, (35), pp. 71-72.

[7] Tao H, He G, Yang J, et al. (2023). MIMU based transmission tower bolt state recognition. *Journal of Vibration and Shock*, 42(20), pp. 98-104. DOI: 10.13465 / j. carolcarrollnki JVS. 2023.20.012.

[8] Liu Y, Chen F, Xie H, et al. (2021). Rapid Detection and positioning Diagnosis of power transmission tower bolts based on Elastic Shock Wave technology. *Electronic Design Engineering*, 29(13), pp. 87-90. DOI:10.14022/ j. ssn1674-6236.2021.13.019.

[9] Zhang A, Deng F (2020). Recognition scheme of external broken vibration of transmission tower based on CNN-RVM. *Computer Simulation*, 37(04), pp. 76-80.

[10] Yang Z, Ao W, Fei X, et al (2020). Research on Intelligent Recognition of transmission tower based on high resolution SAR image and deep learning. *Electrical Measurement and Instrumentation*, 57(04), pp. 71-77. DOI:10.19753/ j. isssn1001-1390.2020.04.012.