

Research Article

Application of Data-Driven Iterative Learning Algorithm in Transmission Line Defect Detection

Yuquan Chen ¹, Hongxing Wang ¹, Jie Shen ¹, Xingwei Zhang ¹ and Xiaowei Gao ²

¹Jiangsu Frontier Electric Technology, Nanjing 211102, China

²Beijing Imperial Image Intelligent Technology, Beijing 100085, China

Correspondence should be addressed to Hongxing Wang; wanghxft@163.com

Received 15 March 2021; Accepted 3 May 2021; Published 13 May 2021

Academic Editor: Shah Nazir

Copyright © 2021 Yuquan Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Deep learning technology has received extensive consideration in recent years, and its application value in target detection is also increasing day by day. In order to accelerate the practical process of deep learning technology in electric transmission line defect detection, the current work used the improved Faster R-CNN algorithm to achieve data-driven iterative training and defect detection functions for typical transmission line defect targets. Based on Faster R-CNN, we proposed an improved network that combines deformable convolution and feature pyramid modules and combined it with a data-driven iterative learning algorithm; it achieves extremely automated and intelligent transmission line defect target detection, forming an intelligent closed-loop image processing. The experimental results show that the increase of the recognition of improved Faster R-CNN network combined with data-driven iterative learning algorithm for the pin defect target is 31.7% more than Faster R-CNN. In the future, the proposed method can quickly improve the accuracy of transmission line defect target detection in a small sample and save manpower. It also provides some theoretical guidance for the practical work of transmission line defect target detection.

1. Introduction

In recent years, deep learning methods produced an effective method for big data processing, and it has made a breakthrough in many different fields such as automatic speech recognition and target recognition [1–5]. At present, it has begun to promote the development of a new generation of artificial intelligence industry worldwide. In the notice of the three-year action plan, China proposed promoting the mutual promotion of the real economy and artificial intelligence technology [6, 7]. With the development of the intelligent industry, deep learning technology has begun to emerge in smart grid image recognition and defect detection applications [8–10].

It is widely known that complex tasks require high-intensity training to construct deep models, and deep learning techniques also require large-scale data for network training [11–13]. In transmission line defect target detection, it is difficult to construct training data

due to the variable target shape. Using a small number of samples will lead to poor generalization of the training model and make it prone to overfitting. These problems bring some difficulties for transmission line defect detection. Combined with the above phenomenon and the important role of recognition algorithm in transmission line defect detection, this paper reviews the related work. For example, The improved SSD method is utilized to detect the transmission line foreign body and present an improved Faster R-CNN deep learning method to accomplish fault detection of the insulator [14, 15]. To solve the low detection accuracy of SSD for the small size object, we proposed an improved algorithm of SSD object detection based on the FP-SSD, which has greatly improved the precision [16]. Chang used SSD combined with binocular vision distance detection method to realize the detection of pantograph offset of transmission line [17]. Antwi-Beko uses Convolutional Neural Network (CNN) to detect and classify defective insulators in transmission line images, achieving high-precision defect

target recognition and location [18]. Wang Yixing trains the stackable automatic encoder (SAE) to initialize and train the deep learning neural network and observes the hidden features of defects from different dimensions, so as to make a preliminary judgment of defects [19]. The above series of works has greatly optimized the target recognition algorithm, but the data is rarely mentioned. The training algorithm designed in this paper provides different ideas to solve the abovementioned problems, mainly using professional knowledge to construct a small amount of labeled data to achieve data-driven iterative training. In order to learn the model incrementally and adapt to unlabeled data, we use the improved Faster RCNN to initialize and train the deep learning neural network and observe the hidden features of defects from different dimensions. In order to make preliminary judgments on the defects, a small amount of labeled data is used to initialize the training network in the early stage of training, and then we continue to perform sample mining in a large amount of unlabeled data in a human-machine collaborative manner [20–22]. Through this data-driven learning method, the recognition module can obtain the recognition model from the training module, and we use the acquired recognition model to label the inspection data, and then the training module uses the labeled raw data to update the model and iteratively improve the model precision. In addition, the deformable convolution and feature pyramid modules are added to the training algorithm to greatly improve the feature extraction capabilities.

2. Materials and Methods

2.1. Training Data. In general, a deep model with good generalization ability must depend on a certain amount of sample data. The richest possible training data including different situations and different forms can make the deep network get more accurate parameters in training. Considering the current lack of publicly available transmission line data sets, it is necessary to construct a professional inspection image insulator defect data set. In order to verify the robustness of the algorithm, a large number of transmission line raw data collected from a certain province and city using UAV inspection are collected. The image resolution is about $5000 * 3000$, covering four seasons. Voltage levels include 35 kV, 110 kV, 220 kV, 500 kV, and high voltage levels. Among them, pin-level defect detection is the most difficult type of defect detection in electrical defect detection, and it is a common electrical defect, so pin defect detection is selected for algorithm testing. To make the network easier to learn, the defects are divided into two categories, including lack of locking pin and rusted-on nut, and a detailed description of the pin-level defect is shown in Figure 1.

2.2. Improved Training Method. Faster R-CNN [23] is the basic two-stage target detection algorithm [24, 25], which has achieved good results in many target detection tasks. Fast

R-CNN is the most basic network architecture to ensure the good scalability of the algorithm. Faster R-CNN introduces Region Proposal Net (RPN) on the basis of Fast R-CNN [26], to participate in the target recognition work, in order to improve the detection speed and realize the integration of the suggestion box generation model and the Fast R-CNN model by sharing the convolutional layer. Considering the complexity of transmission line defect samples, a deformable convolution module and a feature pyramid module are added to the Faster R-CNN algorithm to solve the problem of multiscale target detection to a great extent.

2.2.1. Deformable Convolutional Network. The key point to solve the high-precision recognition of transmission line defects in a complex background is how to adapt to the geometric changes of the proportion, posture, and viewpoint of the defective target in the image. We know that CNN's ability to model geometric transformations is limited, because the geometric structure in CNN is fixed and cannot be deformed during convolution operations. Based on this, a deformable convolution module is introduced to improve the transform modeling ability of CNN. The deformable convolutional network adds a 2-dimensional offset to the sampling point position of the receptive field in the standard convolution. It can freely match the receptive field with the target shape; that is, no matter how the transmission line defect target shape changes, the convolutional receptive field can always cover target range. These offsets are learned from the previous feature map through the additional convolutional layer. Each sampling point has learned to weight and redistribute the modified sampling area, which can achieve more accurate feature extraction, thereby effectively improving the training effect.

The general convolution operations can be divided into two major steps: (1) use a regular grid G on the input feature map for sampling; (2) perform weighting operations. G defines the size and expansion of the receptive field.

The operation of deformable convolution is different. In the regular grid R , it is expanded by adding an offset. For each position P_0 on the output feature map, the calculation equation is as follows:

$$y(p_0) = \sum_{p_n \in R} w(p_n) * f(p_0 + p_n + \Delta p_n), \quad (1)$$

$$G = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}. \quad (2)$$

Among them, f represents the input feature map, y represents the output feature map, and p_n is an enumeration of the positions listed in G . $p_n + \Delta p_n$ represents the offset of the sampling position relative to the center. Since Δp_n is generally not an integer, equation (1) can be realized by the bilinear interpolation method. The specific equation can be expressed as

$$f(p_0 + p_n + \Delta p_n) = \sum_q T(q, p) \cdot f(q), \quad (3)$$

where q represents any position on the input feature map f , and $T(p, q)$ is 0 in most positions of f .



FIGURE 1: Pin-level defect. (a) Rusted-on nut. (b) Lack of locking pin.

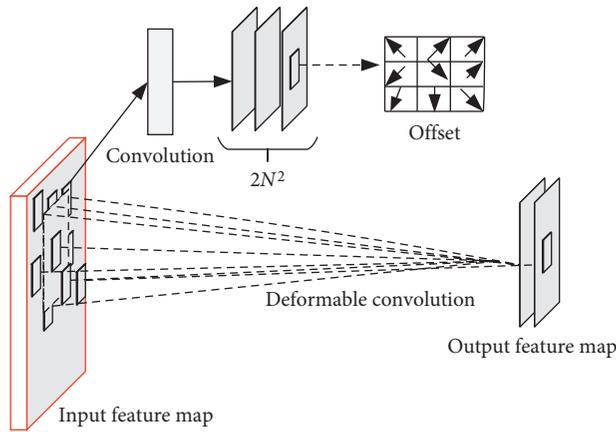


FIGURE 2: Deformable convolution operation.

In Figure 2 in the process of deformable convolution, if the convolution kernel size of the original input feature graph is $N \times N$, use the standard 2D convolution to calculate a new feature map consistent with the input feature map size, and its convolution kernel size is still $N \times N$. The number of channels is $2N^2$ and the $2N^2$ -dimensional vector in the direction of the channel dimension represents N^2 number of 2D offsets, so each position on the new feature map represents the offset of the original convolution kernel on the input feature map. When using the deformable convolution operation, the offset of the corresponding position is superimposed with the sampling position of the original convolution kernel to obtain the offset sampling position. The process after feature sampling is consistent with the standard 2D convolution calculation, and the final result is output feature map.

2.2.2. *Feature Pyramid Network.* The feature pyramid network mainly aims at the recognition of multiscale targets in transmission line defects. By simply changing

the network connection structure, the performance of small target recognition is greatly improved while maintaining the calculation scale of the original model. In traditional vision tasks, multiscale object detection is mainly realized by image pyramid method or hierarchical prediction method, but this method has higher requirements for hardware computing power and memory size, so it can only be used in limited fields. To solve the above problems, feature pyramid network uses the information of each layer in CNN network to generate the final expression feature combination. The feature pyramid network will process the feature output of each CNN layer in the model to generate features that reflect this dimensional information, and there is also an association relationship generated between the features after top-down processing; that is, the high-level features will affect the low-level feature expression. Finally, all the feature combinations are used as the input for the next task of target detection or category analysis. The basic structure of the feature pyramid network is shown in Figure 3. The network is directly modified on the original single

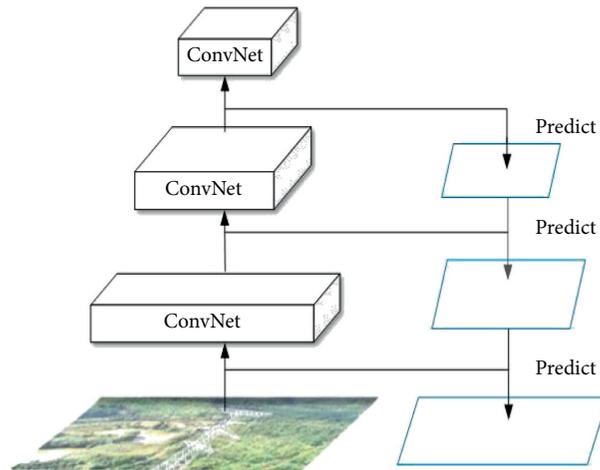


FIGURE 3: Feature pyramid network.

network, and the feature map is sampled from the bottom to the upper layer. The feature map of each resolution is added element by element with the feature map of twice its scale. Using this connection method, the feature maps predicted by each layer are fused with features of different semantic strength and resolution. Targets with different resolutions are detected by feature map with different resolution, so that each layer has moderately strong semantic and resolution features. Moreover, the feature pyramid network only adds additional cross-layer connections on the basis of the original structure. It is confirmed by practical applications that the use of the feature pyramid network hardly increases the amount of extra calculation and time.

3. Data-Driven Training System

In the process of continuous optimization of the model, inspired by the human learning model, a data-driven iterative learning method is constructed. Data-driven iterative learning method mainly solves two problems: one is the automatic construction of in transmission line defect sample library; the other is to solve the problem of overfitting of the training model under small samples.

The data-driven iterative learning method is based on the principle of deep convolutional neural network training. In the early stage, a small number of samples are used for model training. For a large number of unlabeled detection data, the detector with the highest accuracy model is used for detection, and the detection results are sorted by confidence. The threshold is set according to experience, and some of the detection results with higher confidence are extracted, and the active learning method is used for labeling, and the updated data is used to optimize the training model. In this way, by continuously rolling iterative training and detection processes, using self-supervision and active learning methods to cooperate with each other for sample mining, the two processes of training and detection complement each other and continuously improve the accuracy of the recognition model. Mining a large number of unlabeled samples

makes the model not only improve the robustness of the classifier to noise samples or outliers, but also improve the accuracy of detection and finally realize the closed-loop structure of automatic labeling of unlabeled sample data and iterative update of model training, and the closed-loop structure diagram is shown in Figure 4.

How to complete the data connection between the training module and the detection module is a significant step in the application of the data-driven iterative learning algorithm in the transmission line inspection defect detection application. While each process is working on its part, the data dependency between them is resolved through asynchronous calls, so that the two modules can operate at the same time, and they can interact and penetrate each other. Considering the huge amount and variety of transmission line samples, the direct connection of data is mainly through database read and write operations, sample mining is performed in the recognition module part, and the training module reads data for model tuning.

3.1. Data Recognition Module. As the core of data mining, the recognition module provides necessary data support for the training module. The recognition module is based on the deep convolutional neural network model for defect identification, and the detection results are sorted according to the confidence level. If necessary, the human-computer interaction method is used to extract part of the recognition result data with higher confidence, so as to calibrate the type of defects that can be trained and recognized. The identified defect information is stored in the database, and an AI training library for inspection defect data is established. Through the cooperation of self-supervision and active learning methods, the massive uninterrupted image data is structured, labeled, and classified, and training data sources that can be easily found and used were established.

3.2. Data Training Module. With the application of defect recognition and the gradual accumulation of data, the training module can continuously perform feature learning

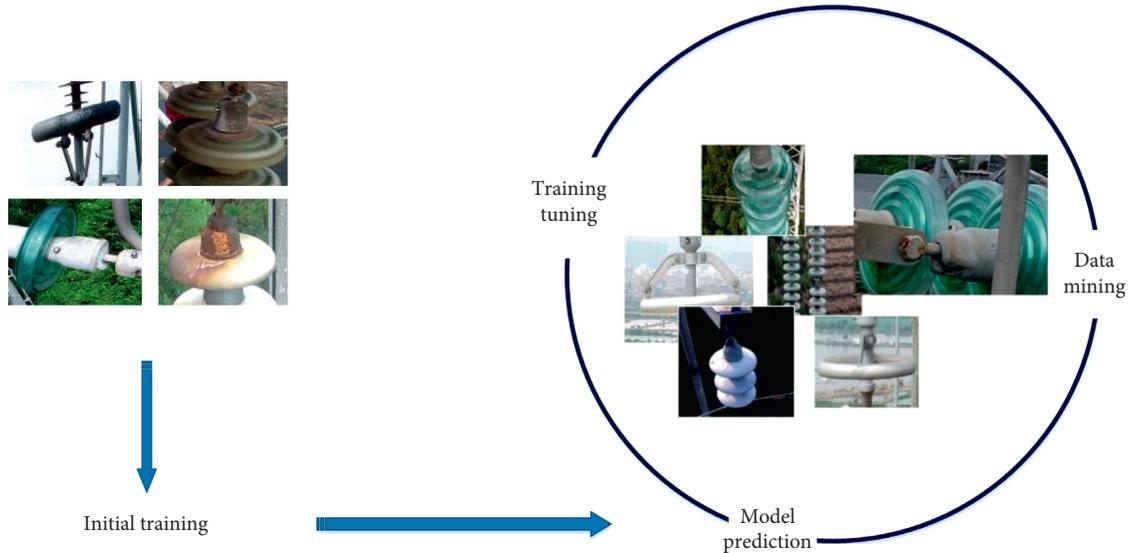


FIGURE 4: Data-driven iterative training closed-loop structure.

and model updates according to predetermined training strategies. It can not only quickly use artificial intelligence technology to improve the efficiency of defect recognition but also customize it according to its own specific needs. Only a small amount of data tags can complete the complex deep learning training task. Data-driven training is realized based on the principle of deep learning, and iterative inspection data and recognition models are continuously rolled over, so as to achieve a closed loop of image processing with high automation and intelligence, and to improve the accuracy of training image defect recognition.

The data-driven iterative learning algorithm ensures the continuous update of the recognition model through the data-driven training function, continuously optimizes the detection effect, and forms a closed-loop ecology from data to training. At the same time, the data-driven training can adaptively adjust training strategies and training modes according to training accuracy without excessive human intervention.

Figure 5 shows the overall process architecture of the data-driven iterative learning algorithm. The underlying environment of the algorithm is Mxnet [27]. The two parts of model recognition and model training are independent of each other. The data-driven iterative learning algorithm can be used to connect the two training and recognition modules in series to train and identify various types of defects independently. The data-driven method is used for sample mining and model iterative updating, and the model with higher accuracy can be selected for publishing.

4. Results and Discussion

Due to the high requirement of hardware configuration for deep learning training. The hardware environment of the algorithm in this experiment is Tesla V100-DGX WORK STATION, the running environment is Ubuntu16.04 operating system, and the computer processor is E5-2698 v4@2.20 GHz. The 101-layer ResNet

[28] is selected as the backbone network, and different initial training parameters are selected in the experiment, such as solver type, initial learning rate, and learning step size. The network optimization function selected in this experiment is Stochastic Gradient Descent (SGD) [29], the learning step is set to 2, the image scaling scale is (1222, 800), the NMS threshold is 0.5, and the initial learning rate is 0.001.

In the initial stage, we made a small number of data samples. The Pascal VOC [30] data set format was directly imitated in the production stage, and we used Extensible Markup Language (XML) to record the location and type of defect target in detail. The experimental data in Table 1 show the initial stage of manually labeling the training set according to the defect characteristics, in which the number of lack of locking pin is 2682, and the number of rusted-on nut data is 1621, and in the test set, the number of lack of locking pin is 1899, and the number of rusts is 1412. Table 1 also shows the growth of data volume after two rounds of data-driven iterative training. Pascal VOC data set analysis uses Average-Precision (AP) as a comprehensive evaluation index and uses precision and recall to examine the model. The above indicators can be expressed as follows:

$$\text{precision} = \frac{TP}{(TP + FP)}, \quad (4)$$

$$\text{recall} = \frac{TP}{(TP + FN)}. \quad (5)$$

In equations (4) and (5), True Positives (TP) represent the number of defective targets that are correctly classified; FP represents the number of background interferences that are mistakenly regarded as defective targets; False Negatives (FN) represents that the defective targets are incorrectly classified as background quantity. AP, as the average accuracy, for the PR curve (Recall on the horizontal axis and Precision on the vertical axis), can be calculated as follows:

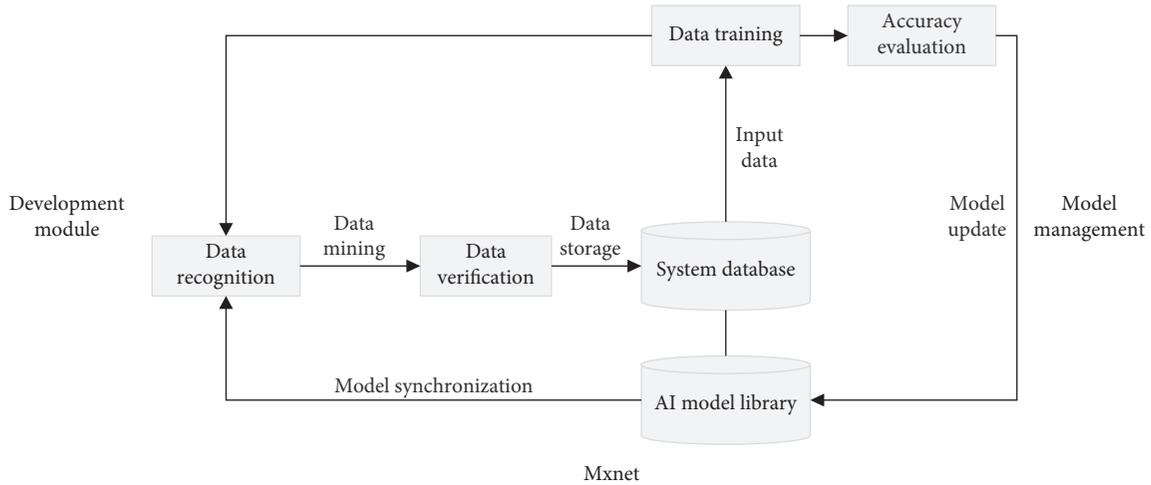


FIGURE 5: The overall architecture of data-driven iterative learning.

TABLE 1: Statistical table of data set quantity.

Datasets	Lack of locking pin (lsqxz)		Rusted-on nut (lmsx)	
	Number of pictures	Number of targets	Number of pictures	Number of targets
Initial training data	2682	3969	1621	2446
The first iterative learning + sample mining	19603	30841	7803	14514
The second iterative learning + sample mining	57272	73415	16238	24840
Test data	1899	2484	1412	1901

$$AP = \int_0^1 p(r)dr. \quad (6)$$

Among them, P and r are, respectively, the accuracy rate and the recall rate. In actual applications, the PR curve is not directly calculated, but the PR curve is smoothed; that is, for each point on the PR curve, the value of precision is the largest precision on the right side of the point.

In order to test the effectiveness of the platform algorithm, this paper designs a comparative experiment, which mainly includes two parts: one is the influence of the setting of the experience threshold of screening data in the process of data-driven training on the accuracy of the model. Based on the improved Faster RCNN algorithm, different experience threshold has been selected to test the accuracy of the model. The second is the analysis of the impact of improved algorithms on model accuracy. To compare the accuracy of training algorithms, first compare the accuracy of the Faster RCNN basic algorithm with the improved Faster RCNN algorithm while keeping the training data consistent, then verify the effectiveness of the data-driven iterative learning algorithm, and finally, test the influence of the number of training iterations on the results of the model.

The first part of the experimental results is shown in Figure 6. The experiment selects five values of 0.65, 0.7, 0.75, 0.8, and 0.85 as the threshold for screening data based on experience. The results show that when the threshold is set to 0.75, the training model obtained by data mining has the highest average precision. After analysis, when the threshold is set higher, some correct data will be screened out, and when the threshold is

lower, it will affect the accuracy of the target. After all training methods tested, obtained when the threshold is about 0.75, the accuracy of the model can be maintained at a high level. Experiments show that the accuracy of the model can be increased to a certain extent by adjusting the appropriate threshold value. Table 2 shows the second part of the experiment content. Compared with Faster R-CNN method, the improved Faster R-CNN method performs more prominently in the detection of lack of locking pin and rusted-on nut defects. At the same time, it can be also seen in Figure 7 that the number of iterations in model training is also an extremely important parameter. As the number of iterations increases, the number of iterations required must increase accordingly to meet the needs of model learning. After the first iterative learning, the model accuracy reached the highest value of 0.704 at 2.06×10^4 iterations. After the second iterative learning, the model accuracy reached the highest value of 0.785 at 3.81×10^4 iterations, which is 25.6% higher than the model accuracy of the initial sample training. In the final experiment, the result shows that the improved Faster R-CNN method has improved the accuracy of the model to a certain extent. But after the second iterative learning, the accuracy is increased by 31.7% compared with Faster R-CNN, which proves that the proposed method is extremely effective.

Finally, the defect images of different scenes are selected to test the model. The detection results are shown in Figure 8. It can be seen that the final model can correctly identify most of the pin defects. In conclusion, the data-driven iterative learning algorithm proposed can effectively improve the accuracy of defect detection and can save manpower to a great extent, which is beneficial to users to independently train high-precision models.

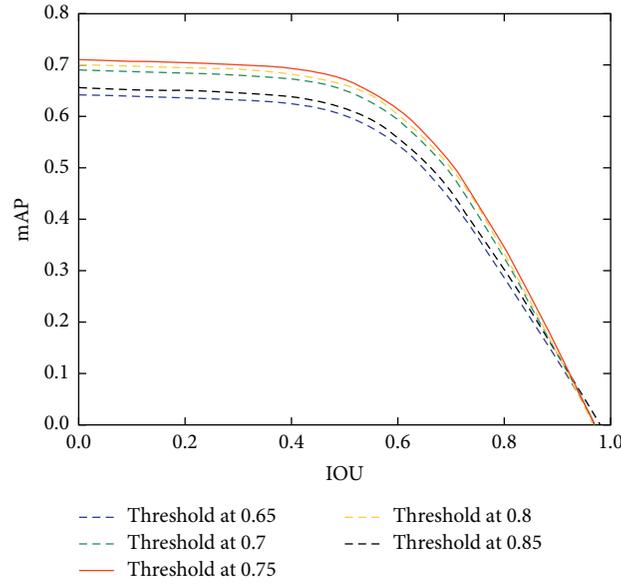


FIGURE 6: Model accuracy under different screening data thresholds.

TABLE 2: The precision of algorithm test.

Algorithm	Classes	Recall	Precision	AP	mAP	Training time (h)
Initial training data + Faster RCNN	lsqxz	0.751	0.684	0.65	0.596	6.11
	lmsx	0.643	0.685	0.541		
Initial training data + feature pyramid + deformable convolution + Faster RCNN	lsqxz	0.78	0.689	0.68	0.625	8.05
	lmsx	0.668	0.702	0.569		
The first iterative learning + sample mining + feature pyramid + deformable convolution + Faster RCNN	lsqxz	0.892	0.794	0.803	0.704	39.05
	lmsx	0.696	0.787	0.604		
The second iterative learning + sample mining + feature pyramid + deformable convolution + Faster RCNN	lsqxz	0.927	0.8	0.854	0.785	118.80
	lmsx	0.752	0.814	0.715		

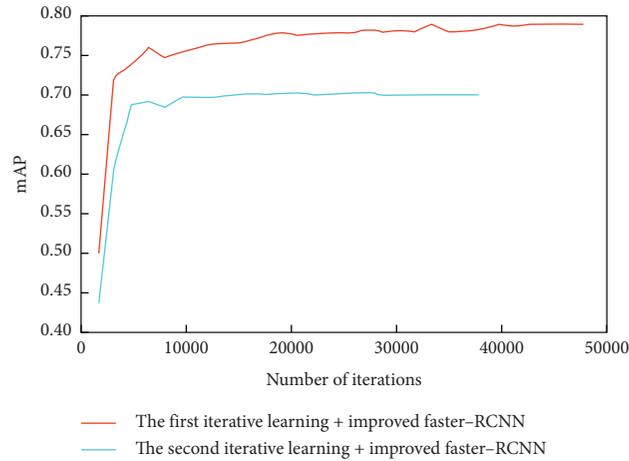


FIGURE 7: Model accuracy display under different iteration times.

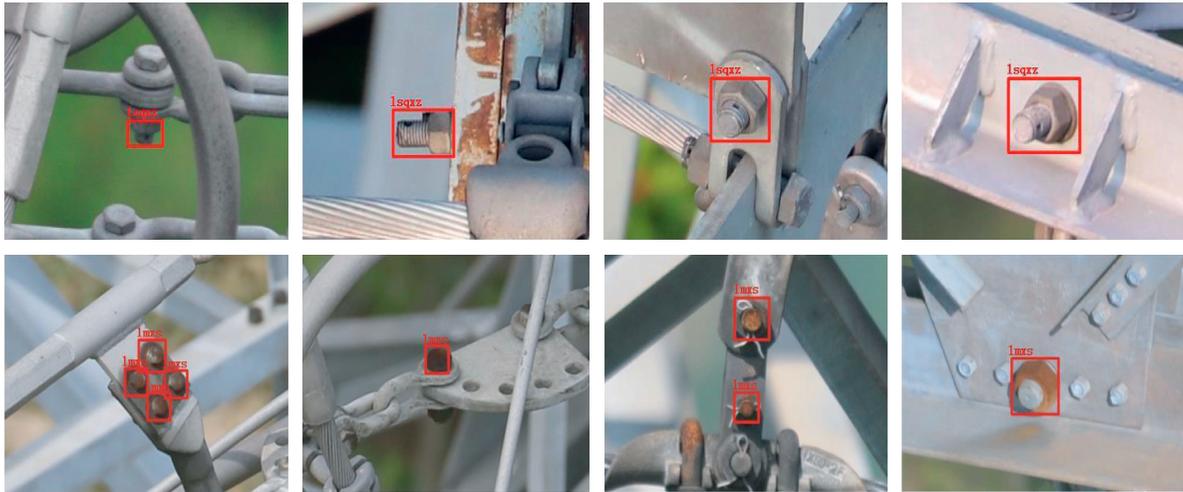


FIGURE 8: Pin defect identification effect.

5. Conclusions

The current work presents a data-driven iterative learning algorithm based on improved Faster R-CNN, which provides various applications such as accelerated training, model detection, and data mining for transmission line defect detection. The data-driven iterative learning algorithm proposes a data-driven training function for the application of transmission line defect detection. After comparison and verification, the proposed method can effectively improve the accuracy of inspection image defect detection. The main conclusions are as follows:

- (1) Faster R-CNN with deformable convolution and feature pyramid modules is chosen as the optimal object detection architecture, so that the value of AP is increased to 0.625. The combined use of deformable convolution and feature pyramid modules makes the improved Faster R-CNN method more prominent in the detection of lack of locking pin and rusted-on nut defects.
- (2) The advantages of data-driven iterative training and improved Faster R-CNN architecture are combined naturally. According to a large number of experimental comparisons, the result showed that the increase of the recognition of data-driven iterative training based on improved Faster R-CNN algorithms for the pin defect target is 31.7% more than Faster R-CNN. The proposed method can achieve a highly automated and intelligent image processing closed loop and improve the accuracy of more transmission line defect target inspections in the future.
- (3) The proposed method provides a means of transmission line data mining. Due to the low cost of data collection and the convenient and efficient algorithm, it has strong practical value, which can meet the needs of users in the transmission line defect detection to a large extent.

Data Availability

This data comes from the original inspection data of the State Grid of China. Due to the limitations of the State Grid, it cannot be used as a public data set.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] L. Du, B. Liu, Y. Wang et al., "Target detection method based on convolutional neural network for SAR image," *Journal of Electronics Information Technology*, vol. 38, pp. 3018–3025, 2016.
- [2] T. Zoughi, M. M. Homayounpour, and M. Deypir, "Adaptive windows multiple deep residual networks for speech recognition," *Expert Systems with Applications*, vol. 139, Article ID 112840, 2020.
- [3] M. Tomaszewski, P. Michalski, and J. Osuchowski, "Evaluation of power insulator detection efficiency with the use of limited training dataset," *Applied Sciences*, vol. 10, no. 6, p. 2104, 2020.
- [4] L. W. Sommer, T. Schuchert, and J. Beyerer, "Deep learning based multi-category object detection in aerial images," in *Proceedings of the SPIE Defense + Security*, International Society for Optics and Photonics, San Francisco, CA, USA, May 2017.
- [5] S. Kamal, S. K. Mohammed, P. R. S. Pillai et al., "Deep learning architectures for underwater target recognition," in *Proceedings of the Ocean Electronics (SYMPOL)*, IEEE, Kochi, India, October 2013.
- [6] G. Ling and L. Xiaohe, "Zhengzhou artificial intelligence industry research," *Innovation Science and Technology*, vol. 36, 2017.
- [7] R. Yang, "The ministry of industry and information technology wants companies to unveil the list and take command of 17 key directions to tackle artificial intelligence," *Computer and Networks*, vol. 44, no. 22, 2018.
- [8] W. Wang, Y. Wang, J. Han, and Y. Liu, "Recognition and drop-off detection of insulator based on aerial image," in

- Proceedings of the 9th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 1, pp. 162–167, Hangzhou, China, December 2016.
- [9] Y. Zhai, D. Wang, M. Zhang, J. Wang, and F. Guo, “Fault detection of insulator based on saliency and adaptive morphology,” *Multimedia Tools and Applications*, vol. 76, no. 9, pp. 12051–12064, 2017.
- [10] U. Tudevtagva, B. Battseren, W. Hardt, and V. Ga-lina, “Image processing based insula-tor fault detection method,” in *Proceedings of the XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)*, pp. 579–583, IEEE, Novosibirsk, Russia, October 2018.
- [11] Y. Liu, B. Lubinski, Y. Shang et al., “Performance comparison of deep learning techniques for recognizing birds in aerial images,” in *Proceedings of the 2018 IEEE Third International Conference on Data science in Cyberspace (DSC)*, IEEE Computer Society, Guangzhou, China, June 2018.
- [12] V. Sharma and R. N. Mir, “A comprehensive and systematic look up into deep learning based object detection techniques: a review,” *Computer Science Review*, vol. 38, 2020.
- [13] A. Hazra, P. Choudhary, and M. S. Singh, “Recent advances in deep learning techniques and its applications: an overview,” in *Proceedings of the ICBEST*, Raipur, India, September 2020.
- [14] J. Yuan, B. Xue, W. Zhang, L. Xu, H. Sun, and J. Zhou, “RPN-FCN based rust detection on power equipment,” *Procedia Computer Science*, vol. 147, pp. 349–353, 2019.
- [15] X. Liu, H. Jiang, J. Chen, and J. Chen, “Insulator detection in aerial images based on faster regions with convolutional neural net-work,” in *Proceedings of the IEEE 14th International Conference on Control and Automation (ICCA)*, pp. 1082–1086, Anchorage, AK, USA, June 2018.
- [16] P. Qin, C. Li, J. Chen, and R. Chai, “Research on improved algorithm of object detection based on feature pyramid,” *Multimedia Tools and Applications*, vol. 78, no. 1, pp. 913–927, 2019.
- [17] L. Chang, Z. Liu, and Y. Shen, “On-line detection of pantograph offset based on deep learning,” in *Proceedings of the 2018 IEEE 3rd Optoelectronics Global Conference (OGC)*, pp. 159–164, IEEE, Shenzhen, China, September 2018.
- [18] E. Antwi-Bekoe, Q. Zhan, X. Xie et al., “Insulator recognition and fault detection using deep learning approach,” *Journal of Physics Conference Series*, vol. 1454, Article ID 012011, 2020.
- [19] Y. Wang, M. Liu, and Z. Bao, “Deep learning neural network for power system fault diagnosis,” in *Proceedings of the 2016, 35th Chinese Control Conference (CCC)*, pp. 6678–6683, IEEE, Chengdu, China, July 2016.
- [20] T. O. Zander and C. Kothe, “Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general,” *Journal of Neural Engineering*, vol. 8, no. 2, Article ID 025005, 2011.
- [21] S. N. Young and J. M. Peschel, “Review of human-machine interfaces for small unmanned systems with robotic manipulators,” *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 99, pp. 1–13, 2020.
- [22] J. Muesing, L. Burks, M. Iuzzolino et al., “Fully bayesian human-machine data fusion for robust dynamic target surveillance and characterization,” in *Proceedings of the AIAA Scitech 2019 Forum*, San Diego, CA, USA, January 2019.
- [23] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [24] L. Yan, M. Yamaguchi, N. Noro, Y. Takara, and F. Ando, “A novel two-stage deep learning-based small-object detection using hyperspectral images,” *Optical Review*, vol. 26, no. 6, pp. 597–606, 2019.
- [25] M. A. Al-Masni, W. R. Kim, Y. Kim et al., “Automated detection of cerebral microbleeds in MR images: a two-stage deep learning approach,” *NeuroImage: Clinical*, vol. 28, 2020.
- [26] R. Girshick, “Fast R-CNN,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448, Santiago, Chile, December 2015.
- [27] T. Chen, M. Li, Y. Li et al., “Mxnet: a flexible and efficient machine learning library for heterogeneous distributed systems,” 2015, <https://arxiv.org/abs/1512.01274>.
- [28] K. He, X. Zhang, S. Ren et al., “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Pittsburgh, PA, USA, June 2016.
- [29] L. Bottou, “Stochastic gradient descent tricks,” *Lecture Notes in Computer Science, Neural Networks: Tricks of the Trade*, Springer, Berlin, Germany, pp. 421–436, 2012.
- [30] S. Vicente, J. Carreira, L. Agapito et al., “Reconstructing pascal VOC,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 41–48, Pittsburgh, PA, USA, June 2014.