ISSN: 2710-0588

Improving the Application of Yolov7 Algorithm in Pin-Losing-Bolts Detection on Power Transmission Lines

Wenqing Zhao, Yutong Zhen*

Computer Science Department, North China Electric Power University, Baoding, 071066, China

*Corresponding author Email: zhenyt0316@163.com

Abstract

A method for bolt deficiency detection on power transmission lines based on YOLOv7 is proposed to address the issues of low detection accuracy and high false negatives in current unmanned aerial vehicle (UAV) inspections. In the Neck module, the existing feature fusion structure is modified by directly combining deep and shallow feature information across adjacent scale feature maps. The prior box sizes in the original YOLOv7 are adjusted to better fit the dimensions and aspect ratios of bolt targets. The improved YOLOv7 algorithm enhances the learning capability of bolt feature information. Experimental results on the bolt deficiency dataset using this method show a 6% increase in precision and a 0.6% increase in mAP_0.5. The experiments demonstrate that the improved method enhances the detection ability of bolt deficiencies on power transmission lines and has practical value in intelligent inspections.

Keywords

Fault Detection; Pin-losing-bolts; YOLOv7; Feature Fusion.

1. Introduction

With the rapid development of China's economic construction and continuous improvement of infrastructure, the demand for electricity in cities has also increased. Consequently, the power system is experiencing a growing load, leading to higher requirements for its safety and stable operation. As a result, the workload of daily inspections for the power system has increased, with a particular emphasis on the inspection of power transmission lines. Most of these transmission lines are located in outdoor areas far from cities with unfavorable environmental conditions.

Previously, inspections heavily relied on manual labor, where inspectors used their eyes and handheld devices. However, this inspection method is not only costly and inconvenient in terms of transportation but also poses high risks[1]. Moreover, issues like line-of-sight obstructions and human error impact the quality and efficiency of inspections.

In recent years, unmanned aerial vehicle (UAV) technology has undergone rapid development and has also revolutionized the way power grid inspections are conducted. Remote-controlled UAV inspections are gradually replacing traditional manual inspections. UAVs can autonomously plan inspection routes using systems like Beidou and GPS, enabling automated inspections. Equipped with high-definition cameras, UAVs capture images of the power transmission lines' basic condition and upload them for analysis. Remote personnel can then remotely view real-time videos or images to conduct initial diagnostics of the power transmission lines.Compared to traditional field inspections conducted by personnel, remote UAV inspections are more efficient, convenient, and safe. They significantly reduce labor and maintenance costs[2].

ISSN: 2710-0588

Defect detection in power transmission lines primarily focuses on critical components that are susceptible to damage, such as insulators, fittings, and bolts. Among these components, bolts play a fundamental role in fastening and connecting various parts of the power transmission lines and are the most abundant in quantity. As the power transmission lines are mostly exposed to outdoor environmental conditions, they are susceptible to damage caused by natural weather elements. Damaged bolts can pose significant safety hazards, including line loosening. Therefore, timely detection and repair of bolt deficiencies in power transmission lines have always been a challenging problem in power inspections.

Compared to targets such as insulators and fittings, bolts have a smaller size. When unmanned aerial vehicles (UAVs) capture images in complex terrains, bolts in the aerial images of power transmission lines typically occupy a small proportion. Due to their small size and the complex background of the images, bolt deficiency detection is a challenging task in UAV inspections. The detection of bolt deficiencies is highly susceptible to issues such as image quality, making it a persistent challenge in UAV inspections.

In recent years, deep learning technology has made significant advancements in the field of object detection. Object detection algorithms such as SSD[3] (Single Shot MultiBox Detector), Faster R-CNN[4], and YOLO[5-6] (You Only Look Once) have been extensively validated and applied in power inspections.

Li Xuefeng[7] proposed an improvement to the Faster R-CNN algorithm called PinNet, which further enhances the semantic and positional information at the lower levels, thereby improving the ability to detect small targets[8]. Xue Yang [9-10]proposed a method for bolt detection in aerial images by first performing data augmentation to expand the dataset and then training a convolutional neural network-based object detection model. The acquired image data was augmented through techniques such as flipping, translation, and rotation to increase the dataset size. The Faster R-CNN model with ResNet as the backbone network was used for training and testing[11-13]. Experimental results demonstrated that this method achieved high detection accuracy for bolts on piercing line clamps.

This paper addresses the issues in bolt deficiency detection in power transmission lines and proposes improvements by applying YOLOv7, which demonstrates good accuracy and speed performance, to the detection of bolt deficiencies. Given the small size of bolt targets, the following modifications were made: In the Neck part, the information transfer path was altered. The different-scale feature maps extracted by the Backbone, containing rich information, were fused more effectively to avoid significant feature loss during multiple fusion steps. This enhances the network's capability to extract features from small targets. Lastly, the prior boxes in the network were optimized and adjusted to better suit the size of bolt targets[14-15].

2. Object Detection Method based on YOLOv7

In 2016, Joseph Redmon introduced the original YOLO (You Only Look Once) algorithm. Over the following years, various iterations and improvements were made, resulting in algorithms such as YOLOv1, YOLOv3, YOLOv4, and so on. YOLOv7 was released in 2022 and further divided into versions like YOLOv7-X, YOLOv7-W6, YOLOv7-E6, based on network depth and width.

The YOLOv7 architecture consists of three main components: Backbone, Neck, and Head. The Backbone is composed of modules like CBS, ELAN, and MP-1, which serve as the backbone network for feature extraction using multiple convolutional layers. The Neck performs feature fusion using a combination of "FPN (Feature Pyramid Networks) and PAN (Path Aggregation Network) " methods.

The Head comprises modules such as CBS, SPPCSPC, E-ELAN, MP-2, and RepConv, and its main function is feature fusion to enrich the feature information of different-sized feature maps. The Head is divided into three or four detection heads depending on the configuration file,

ISSN: 2710-0588

responsible for final classification and regression on different-sized feature maps. It outputs classification information, background/foreground judgments, and coordinate information[16-18].

3. Improvement and Optimization of YOLOv7 Model

This paper builds upon the YOLOv7 algorithm and introduces detailed improvements to achieve more accurate localization and higher discrimination rates in bolt deficiency data. The improvements primarily focus on the Backbone and Neck parts of the model. Fig. 1 illustrates the modified structure of YOLOv7.

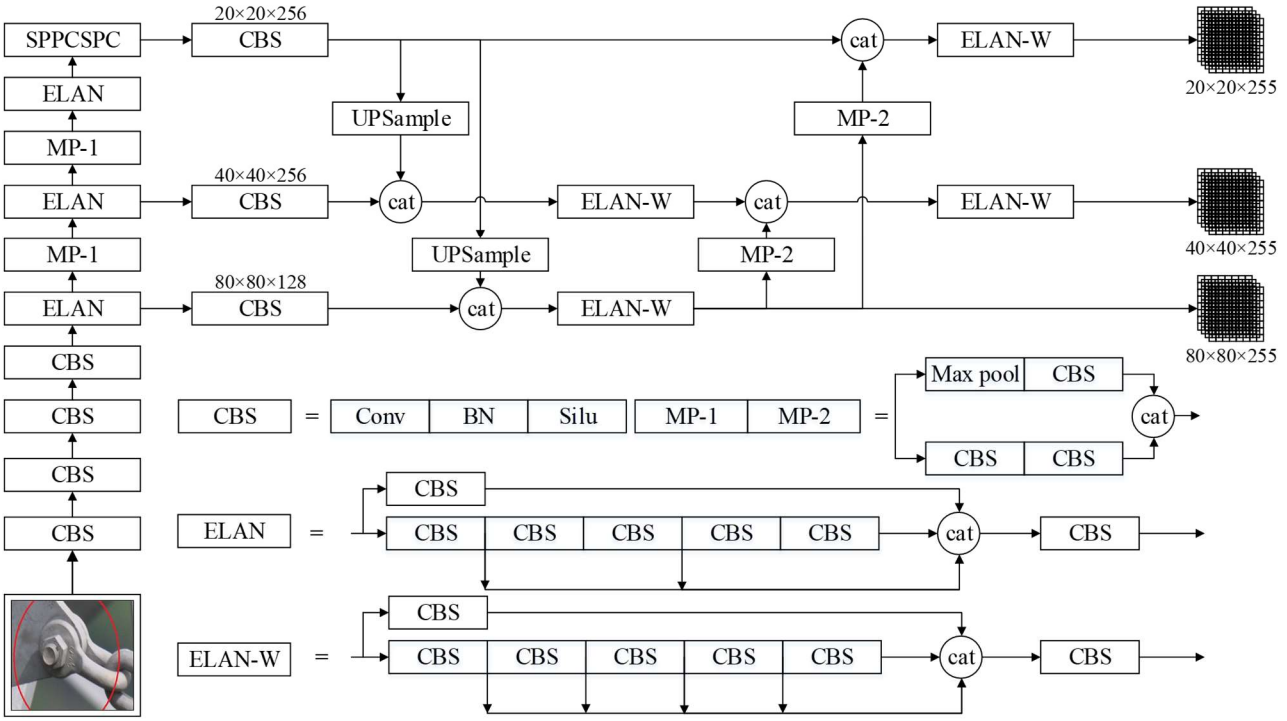

Fig.1 Improved YOLOv7 structure diagram

3.1. Building CL-FPN Structure

Bolt targets in regular unmanned aerial vehicle (UAV) inspection aerial images have a smaller proportion compared to larger components such as fittings and insulators. Most deep learning models are unable to extract detailed features of bolt targets during the training process. In the Neck part of the YOLOv7 model, a feature fusion module is employed to enrich the semantic information obtained from deep networks. By fusing the low-resolution feature maps and the high-resolution feature maps with less semantic information from shallow networks, the model improves its discrimination accuracy for small-sized targets.

The YOLOv7 model adopts two information propagation paths, namely the "top-down" Feature Pyramid Network (FPN) structure and the "bottom-up" Path Aggregation Network (PAN) structure, for feature fusion of different-sized feature maps. In traditional FPN, high-level feature maps need to undergo multiple upsampling and feature fusion with feature maps from adjacent paths before being propagated to the bottom-level network. This information propagation path may lead to the gradual loss of rich semantic information possessed by high-level feature maps through multiple upsampling and feature fusion, preventing the bottom-level feature maps from learning the rich semantic information from the high-level feature maps.

ISSN: 2710-0588

To address this issue, this paper optimizes the feature fusion module of the original YOLOv7 model and proposes Spanning Connections-Feature Pyramid Network (CL-FPN). CL-FPN directly fuses the upsampled high-level feature maps with the bottom-level feature maps, avoiding the loss of semantic information through multiple fusion steps. By directly fusing the high-level feature maps with the bottom-level feature maps, CL-FPN can enhance the resolution of the feature maps while enriching semantic information, thereby improving the model's detection ability for small targets. Fig. 2 illustrates the modified Backbone and Neck structures with the addition of CL-FPN.

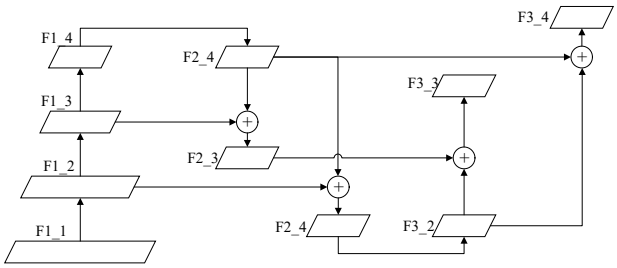

Fig.2 Backbone and Neck structures after adding CL-FPN

3.2. Spanning Connections-Feature Pyramid Network are as follows:

(1) The input image of size 640×640×3 undergoes multiple convolutions and one spatial pyramid pooling in the Backbone. This process results in feature maps with the following abstract sizes: F1_1 of size 160×160×128, F1_2 of size 80×80×256, F1_3 of size 40×40×512, and F1_4 of size 20×20×1024.F1_4 is then subjected to convolution, normalization, and dimension reduction to obtain F2_4. After upsampling, F2_4 is concatenated with F1_3 along the dimension. Subsequently, a convolution operation is performed to adjust the channel number, resulting in a feature map of size 40×40×1024, denoted as F2_3.At the same time, F1_4 undergoes convolution, normalization, and four-fold upsampling. It is then concatenated with F1_1 along the dimension. Following another convolution operation to adjust the channel number, a feature map of size 80×80×1024, denoted as F2_2.

(2) F2_2 is subjected to another convolution operation, resulting in a feature map of size 80×80×256, denoted as F3_2. F3_2 continues downsampling to propagate information upwards. It is then concatenated with the feature map F2_3, resulting in a feature map of size 40×40×512, denoted as F3_3.Simultaneously, F2_2 is downsampled, resulting in a feature map of size 20×20×256. This downsampled feature map is concatenated with F2_4, and after convolution and normalization, a feature map of size 20×20×1024, denoted as F3_4.

(3) Finally, F3_2, F3_3, and F3_4 are passed into the Head module for the final classification and regression of different-sized feature maps. The output includes classification information, determination of background and foreground, as well as coordinate information.

3.3. Optimization Method for Anchor Box setting based on K-media++

YOLOv7 is a single-stage object detection algorithm based on anchor boxes. Like other anchor box based object detection algorithms, YOLOv7 sets dense anchor boxes on the image and trains a network to predict the classification confidence and position offset of the anchor boxes. Among them, the parameters of the anchor box as the initial setting of the algorithm are an important factor affecting the training effect of the YOLOv7 algorithm.

The overall optimization process is based on the k-media++clustering algorithm. K-media++is an improved algorithm for k-means, which mainly uses the median of samples instead of the mean as the basis for generating new cluster centers, reducing the impact of noisy samples on cluster center selection and improving algorithm robustness.

The specific process of optimizing anchor box parameters is as follows:

ISSN: 2710-0588

(1) Use the k-media++algorithm to initialize the anchor box parameters, and use the YOLOv7 algorithm to train the model under the initial anchor box until the network converges.

(2) On the basis of the convergence model, the target objects matched by the last three detection layers are extracted separately through the anchor box matching stage.

(3) Cluster the targets matched by three different sized detection layers separately, and the obtained cluster centers are set as new anchor box parameters corresponding to the three different sized detection layers to complete optimization.

4. Experimental Results and Analysis

4.1. Experimental Data

A total of 464 images of transmission line bolt defects were collected in this study. Data cleaning was performed first, which involved cropping excessively large images and removing low-quality, highly similar, and small-sized images. After the data cleaning process, 428 usable images of transmission line bolts remained for further detection. These images were clear in quality and had suitable sizes. They were annotated using LabelImg according to the "Annotation Specification for Defect Images of Overhead Transmission Line Equipment (Trial)" and saved in YOLO format. The labels for the bolt samples were divided into two categories: "visible-pin-losing" (missing bolt) and "normal bolt." This resulted in a preliminary dataset of bolt images suitable for experimentation.

Statistical analysis revealed that the average area ratios of the two categories of bolts in the original images were 0.034 and 0.023, respectively. It is generally considered that objects with an area ratio of less than 0.1 in the original image are classified as small objects in object detection, which makes their detection more challenging.

Table 1. The number of bolts in the dataset and the area ratio of the original image

	missing bolt	normal bolt
Number of training sets	2486	2492
Number of validation sets	367	299
Number of test sets	298	301
Target average original area proportion	0.034	0.023

To enhance the diversity of the data and ensure the generalization capability of the proposed method, data augmentation was performed in this study. The following augmentations were applied to the original images:

Brightness Enhancement: The brightness of the images was increased by 30% and decreased by 30% to create variations in lighting conditions.

Rotation and Flipping: The images were rotated by 90 degrees and flipped horizontally to introduce variations in orientation.

Using the above augmentation techniques, a total of 2,395 bolt images and their corresponding annotated files were generated. To conduct the experiments in this study, the generated data was split into training, validation, and testing sets in an 8:1:1 ratio. The statistical results of the dataset are summarized in Table 1.

4.2. Experimental Environment and Parameters

The experiment was conducted using Python programming language. The GPU used was NVIDIA GeForce GTX 3070Ti, and the deep learning framework used was PyTorch 1.13.1, with CUDA version 12.3.

ISSN: 2710-0588

During the network training phase, the input image size was set to 640×640. The batch size was set to 16, and the number of training epochs was set to 200. The optimizer chosen was SGD (Stochastic Gradient Descent). The initial learning rate was set to 0.2. During the initial training, the Warmup method was used to gradually increase the learning rate. Once the model became stable, the learning rate was adjusted to a certain value. After that, a cosine annealing strategy was used to adjust the learning rate.

4.3. Ablation Experiment

In this experiment, the evaluation metrics used were Precision (P), Recall (R), Average Precision (AP), and mean Average Precision (mAP_0.5) with an IoU threshold of 0.5.

To validate the effectiveness of the proposed method, the paper conducted ablation experiments on the improvements made to YOLOv7. The experimental results are presented in Table 2. The proposed method refers to the combination of the two aforementioned improvements in the YOLOv7 algorithm.

Table 2. Ablation experiment results

model	P	R	mAP_0.5	mAP_0.5: 0.95
YOLOv7	0.922	0.744	0.814	0.619
YOLOv7+CL-FPN	0.939	0.738	0.809	0.650
YOLOv7+ K-media++	0.920	0.754	0.807	0.657
proposed method	**0.950**	**0.745**	**0.823**	**0.652**

From the above table, it can be seen that the detection accuracy of the method in this article for missing bolts reaches 95%, with high accuracy. Compared to the previous YOLOv7, it has improved accuracy by 2.8% while maintaining a similar recall rate, reducing missed and false detections; The detection effect has been greatly improved.

This article will sequentially incorporate the improvement points mentioned above into YOLOv7, resulting in varying degrees of improvement in detection accuracy. After adding CL-FPN, the detection accuracy of missing bolts increased by 1.7%. Because the area of bolt targets in the original image is relatively small, the detection of small targets relies more on fine-grained information in shallow networks. Cross layer feature fusion can fully integrate the detailed information in deep networks and the positional information in shallow networks, enabling the network to learn richer bolt features during the training process. After optimizing the anchor box parameters, the recall rate of missing bolts increased by 1%. This is because the optimized anchor box parameters are more suitable for bolt missing data sets, making the network more accurate in locating bolt targets.

4.4. Comparative Experiment

To further validate the effectiveness of the method proposed in this article, experiments were conducted on the dataset using current mainstream object detection models. To comprehensively consider the performance of different algorithms. The experimental results are shown in Table 3.

In terms of accuracy, YOLOv7 has the highest accuracy among Faster RCNN, YOLOv3, YOLOv5, and YOLOv7, with an accuracy 2.8% higher in this article. Compared with other methods, our method has significant advantages in accuracy, recall, precision, and average precision.

Due to the fact that Faster RCNN belongs to a two-stage object detection algorithm, which first judges the foreground and background during the detection process, and then classifies them, it has an advantage in terms of recall rate; The YOLO series belongs to a single-stage object

ISSN: 2710-0588

detection algorithm, which completes classification and regression simultaneously. Therefore, YOLOv7 and the method proposed in this article have fast speed and high accuracy.

Compared with other mainstream object detection algorithms, the accuracy is the highest and the recall rate is also the highest among single-stage object detection algorithms, so overall, the performance of our method is superior. It has certain application value in the detection of bolt shortage in transmission lines.

Table 3. Comparison of experimental results between different algorithms

model	P	R	mAP_0.5	mAP_0.5: 0.95
Faster-RCNN	0.574	0.832	0.756	0.492
YOLOv3	0.782	0.697	0.72	0.464
YOLOv5	0.83	0.738	0.785	0.537
YOLOv7	0.922	0.744	0.814	0.619
proposed method	**0.950**	**0.745**	**0.823**	**0.652**

5. Conclusion

This article proposes a bolt shortage detection method for transmission lines based on YOLOv7. Improvements have been made to YOLOv7: CL-FPN structure has been constructed in the Neck section, which fully combines the rich localization information obtained from shallow networks with the rich semantic information obtained from deep networks; After optimizing the anchor frame parameters, the model is more suitable for bolt shortage and sales datasets. The experiment shows that the improved YOLOv7 algorithm improves the detection accuracy, reduces missed detections, has stronger positioning ability, and meets the requirements of real-time detection in the bolt shortage and sales dataset. It provides a new reference scheme for bolt shortage and sales detection in intelligent power grid inspection.

References

[1] YANG Hao, LIN Nan, Yuan Chen, et al. Cyber security defense of unmanned aerial vehicle in power utilities[J]. Journal of electric power science and technology, 2021,36(04):172-180 (in Chinese).

[2] ZHAO Zhenbing, JIANG Zhigang, LI Yanxu, et al. Overview of visual defect detection of transmission line components[J]. Journal of image and graph, 2021,26(11):2545-2560 (in Chinese).

[3] Wei Liu, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]. // Proceedings of the 14th Europen Conference on Computer Vision. Amsterdam, Netherlands,2016:21-37.

[4] Shaoqing Ren, Kaiming He, Ross Girshick, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017,39(6):1137-1149.

[5] ZHANG Shu, WANG Haotian, DONG Yaochong, et al. Bolt detection technology of transmission lines based on deep learning[J]. Power system technology, 2021,45(07):2821-2829 (in Chinese).

[6] Redmon J, Divvala s, Girshick R, et al. You only look once: unified, real time object detection[C].// IEEE Conference on Computer Vision and Pattern Recognition. LasVegas, USA, 2016: 779-788.

[7] Li Xuefeng, Liu Haiying, Liu Gaohua, Su Hansong. Transmission Line Pin Defect Detection Based on Deep Learning [J/OL]. Power Grid Technology :1-9[2020-11-03].https://doi.org/10.13335/j.1000-3673.pst.2020.1028. (in Chinese).

[8] Redmon J, Farhadi A .YOLO9000: better, faster, stronger[C].//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu,USA, 2017: 6517-6525.

[9] Xue Yang, Wu Haidong, Zhang Ning, etc Based on the improved Faster R-CNN transmission line puncture clamp and bolt detection [J] Progress in Laser and Optoelectronics, 2020, 57 (8): 8(in Chinese).

ISSN: 2710-0588

[10] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection [EB/OL].(2020-06-22). https://arxiv.org/abs/1804.02767.

[11] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, USA 2017: 2117-2125.

[12] Shu Liu, Lu Qi, Haifang Qin, et al. Path Aggregation Network for Instance Segmentation[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City, USA,2018: 8759-68.

[13] ZHAI Yongjie, YANG Xu, ZHAO Zhenbing, et al. Integrating co-occurrence reasoning for Faster R-CNN transmission line fitting detection[J]. CAAI transactions on intelligent systems, 2021, 16(2): 237–246. (in Chinese).

[14] ZHAO Wenqing, CHENG Xingfu, ZHAO Zhenbing, et al. Insulator recognition based on attention mechanism andFaster RCNN[J]. CAAI transactions on intelligent systems, 2020, 15(1): 92–98 (in Chinese).

[15] Mingxing Tan, Le Q.V. Rethinking model scaling for convolutional neural networks[C]//36th International Conference on Machine Learning. Los Angeles, USA, 2019:6105-6114.

[16] ZHAO Yongqiang, RAO Yuan, DONG Shipeng, et al. Survey on deep learning object detection[J]. Journal of image and graphics,2020,25(04):629-654 (in Chinese).

[17] TANG Hong, FAN Sen, Tang Fan, et al. Recommendation algorithm combining knowledge graph and attention mechanism[J]. Computer engineering and applications, 2022,58(05):94-103 (in Chinese).

[18] ZHAO Yongqiang, RAO Yuan, DONG Shipeng, et al. Survey on deep learning object detection[J]. Journal of image and graphics,2020,25(04):629-654 (in Chinese).